Kansanterveyslaitos
Folkhälsoinstitutet
National Public Health Institute

Publications of the National Public Health Institute A 1 / 2004

Jani Saarela

CHARACTERIZATION OF ASPARTYLGLUCOSAMINIDASE ACTIVATION AND ASPARTYLGLUCOSAMINURIA MUTATIONS

Department of Molecular Medicine,
National Public Health Institute, Helsinki, Finland and
Department of Medical Genetics,
University of Helsinki, Finland

Helsinki 2004

CHARACTERIZATION OF ASPARTYLGLUCOSAMINIDASE ACTIVATION AND ASPARTYLGLUCOSAMINURIA MUTATIONS

Jani Saarela

Department of Molecular Medicine,
National Public Health Institute, Helsinki, Finland and
Department of Medical Genetics,
University of Helsinki, Finland

Academic Dissertation

To be publicly discussed with the permission of the Medical Faculty of the University of Helsinki, in the lecture room 3 of Biomedicum Helsinki, Haartmaninkatu 8, Helsinki, on January 30th, 2004, at 12 o'clock noon.

Helsinki 2004

Supervised by

Professor Leena Peltonen-Palotie
National Public Health Institute
and Department of Medical Genetics
University of Helsinki, Helsinki, Finland

Reviewed by

Professor Ole Kristian Tollersrud
Department of Medical Biochemistry
University of Tromsoe
Tromsoe, Norway

and

Docent Marc Baumann
Protein Chemistry/Proteomics Unit
and Neuroscience Research Program
University of Helsinki, Helsinki, Finland

To be publicly discussed with

Professor Marja Makarow
Institute of Biotechnology
and Department of Applied Biochemistry and Molecular Biology
University of Helsinki, Helsinki, Finland

Julkaisija-Utgivare-Publisher

Kansanterveyslaitos (KTL)
Mannerheimintie 166
00300 Helsinki
puh. vaihde 09-47441, felefax 09-4744 8408

Folkhälsoinstitutet
Mannerheimvägen 166
00300, Helsinki
tel. växel 09-47441, telefax 09-4744 8408

National Public Health Institute
Mannerheimintie 166
00300 Helsinki, Finland
phone +358-9-47441, telefax +358-9-4744 8408

Publications of the National Public Health Institute, KTL A1/2004

ISBN 951-740-407-7
ISSN 0359-3584
ISBN 951-740-408-5 (PDF version)
ISSN 1458-6290 (PDF version)
http://ethesis.helsinki.fi

Cosmoprint Oy
Helsinki 2004

To my family

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original articles, which are referred to in the text by their Roman numerals:

I. **Saarela, J.**, Laine, M., Tikkanen, R., Oinonen, C., Jalanko, A., Rouvinen, J. and Peltonen, L. (1998) Activation and oligomerization of aspartylglucosaminidase. *J. Biol. Chem.*, **273**, 25320-8. *

II. **Saarela, J.**, Laine, M., Oinonen, C., von Schantz, C., Jalanko, A., Rouvinen, J. and Peltonen, L. (2001) Molecular pathogenesis of a disease: structural consequences of aspartylglucosaminuria mutations. *Hum. Mol. Genet.*, **10**, 983-95.

III. **Saarela, J.**, Oinonen, C., Jalanko, A., Rouvinen, J., and Peltonen, L. (2003) Autoproteolytic activation of human aspartylglucosaminidase. In press, The Biochemical Journal.

In addition, some unpublished data are presented.

* Appeared also in the thesis of Carita Oinonen (2000).

ABBREVIATIONS

AADG	2-acetamido-1-β-(L-aspartamido)-1,2-dideoxy-D-glucose
AGA	aspartylglucosaminidase
AGU	aspartylglucosaminuria
AGU_{Fin} major	double mutation 482G->A + 488G->C, R161Q + C163S
bp	base pair
CA	cephalosporin acylase
CD-MPR	cation-dependent mannose 6-phosphate receptor
cDNA	complementary deoxyribonucleic acid
CI-MPR	cation-independent mannose 6-phosphate receptor
COS-1	African green monkey kidney cells
del	deletion
DNA	deoxyribonucleic acid
ER	endoplasmic reticulum
FPLC	fast performance liquid chromatography
fs	frame shift
GAT	glutamine amidotransferase
GCA	glutaryl 7-aminocephalosporanic acid acylase
GlcNAc	N-acetylglucosamine
hsp	heat-shock protein
ins	insertion
kb	kilobase pair
kDa	kilodalton
K_M	Michaelis constant
MPR	mannose 6-phosphate receptor
mRNA	messenger ribonucleic acid
Ntn	N-terminal nucleophile
PAGE	polyacrylamide gel electrophoresis
PDI	protein disulfide isomerase
PGA	penicillin G acylase
PPIase	peptidyl-prolyl isomerase
PRO	proteasome β-subunit
PRPP	5-phosphoribosyl-1-pyrophosphate
rmsd	root mean square deviation
SDS	sodium dodecyl sulphate
TGN	*trans*-Golgi network
V_{max}	maximal reaction velocity
WT	wild type
Å	Ångström (10^{-10} m)

In addition, standard one-letter abbreviations of nucleic acids and amino acids and three-letter abbreviations of amino acids are used.

ABSTRACT

Lysosomal aspartylglucosaminidase (AGA) participates in glycoprotein breakdown by catalyzing cleavage of the N-glycosidic bond between asparagine and N-acetylglucosamine. Mutations in the *AGA* gene lead to a lysosomal storage disease, aspartylglucosaminuria (AGU), which is characterized by progressive mental retardation of patients. The recessively inherited disease is enriched in the Finnish population.

Prior to this study, the three-dimensional crystal structure and the catalytic mechanism of human AGA had been determined (Oinonen *et al.* 1995, Tikkanen *et al.* 1996a). For the activation of AGA, dimerization of AGA precursor polypeptides had been proposed to be required (Riikonen *et al.* 1996). However, detailed knowledge of the initial steps of the maturation of AGA was however lacking. This study was carried out to characterize the dimerization process and the autocatalytic activation mechanism of AGA in detail.

The native AGA molecule consists of two αβ-dimers in close contact. The dimer interface is rather flat and the dimers are linked by non-covalent interactions. On the dimer interface, three regions of amino acids were found to provide an abundance of contacts between the dimers. Mutagenesis of these regions prevented dimerization and activation of precursor molecules.

Site-directed mutagenesis at the activation site revealed that the side chain hydroxyl group of T206 is required for activation of the AGA precursor polypeptides. In addition, correct conformation of the polypeptide backbone between the scissile D205-T206 peptide bond is crucial for triggering the activation, using the energy contained in a high-energy strained peptide bond. Other active site amino acids also contribute to the activation by forming a structural scaffold which stabilizes the residues directly involved in autocatalysis. On the basis of these studies, a reaction mechanism for the activation of AGA was proposed.

Seven novel AGU mutations were found, including mutations in three Finnish and four non-Finnish patients. These and the majority of the previously found AGU mutations were expressed *in vitro* and their consequences at the polypeptide structural level were characterized in detail. These data will help in dissecting the molecular pathogenesis of the AGU disease.

Finally, the expression and purification of the recombinant human AGA precursor has provided material for crystallization experiments, which hopefully will lead to determination of the structure of the human AGA precursor molecule.

REVIEW OF THE LITERATURE

1. Lysosomes

Lysosomes, discovered in 1955, are the terminal degradative compartment of the endocytic pathway and are also involved in phagocytosis and autophagy (de Duve *et al.* 1955, de Duve 1983, Kornfeld and Mellman 1989). These cytoplasmic acidic vesicles surrounded by a single membrane are rich in hydrolytic enzymes, which are used for the controlled intracellular digestion of macromolecules (Novikoff 1963). Acidic environment of pH 4.6-5.0 is maintained by H^+ ATPase at the lysosomal membrane, and membrane potential is dissipated by influx of Cl^- and efflux of K^+ and Na^+, providing an optimal environment for at least 50 types of acid hydrolases in the lysosomal lumen (Bainton 1981, Mellman *et al.* 1986, Stevens and Forgac 1997). Lysosomes are very diverse in size and shape and their identification can be performed by histochemical visualization of acid phosphatase (de Duve 1963). The diversity reflects the wide variety of digestive functions mediated by acid hydrolases – the breakdown of intracellular and endocytosed debris, the destruction of phagocytosed microorganisms, and the production of nutrients for the cell. The resulting free amino acids, lipids, and oligosaccharides can be used as substrates for synthesis of new macromolecules or for degradation as sources of energy. The action of lysosomal hydrolases finally results in complete degradation of macromolecules into monomeric residues. Specific transporters in the lysosomal membrane move them out of the lysosomes to be metabolically reused (Saier 1999). Recently, a better understanding of the endocytic pathway has challenged the function of lysosomes. The current hypothesis indicates that lysosomes, defined by the absence of mannose 6-phosphate receptors, are actually storage granules for acid hydrolases. The digestion of endocytosed macromolecules is proposed to occur in hybrid vesicles that are formed by fusion of lysosomes and late endosomes (Luzio *et al.* 2000). Additionally, in many cell types lysosomes are capable of fusion with the plasma membrane to secrete their contents (Stinchcombe and Griffiths 1999).

2. Protein folding, modification, and degradation

Proteins entering the secretory pathway – including proteins of the endoplasmic reticulum, Golgi, endosomes/lysosomes, nuclear and plasma membranes, and extracellular space - are synthesized on the cytoplasmic side of the endoplasmic reticulum (ER) membrane in ribosomes. Recognized by their signal sequence, these newly synthesized proteins are translocated into the lumen of the ER through the translocon complex, a proteinaceous channel through which the emerging chain enters the ER lumen or integrates with the ER membrane. The lumen of the ER contains molecular chaperones and folding factors that

assist protein folding in all stages. In the ER, several co-translational and post-translational modifications take place, which are important for correct protein folding and transport.

2.1. Protein folding

Protein folding starts co-translationally as the nascent proteins are being synthesized at the ribosomes. Folding typically occurs in three stages. First, co-translational and co-translocational protein folding occurs in the translocon complex, a membrane protein complex utilized to insert the nascent polypeptide chain into the membrane and transfer soluble polypeptides across the lipid bilayer. Early stages of protein folding proceed quickly, for secondary structure formation and compaction of the structure require much less than one second (Roder and Colon 1997). Post-translational folding takes place after the completed chain has been released from the ribosome and the translocon complex. The resulting molten globules are more open and less ordered than the final folded form of the protein (Feng *et al.* 1994). Finally, when the subunits have reached a conformation that is close to the final folded state, oligomeric assembly takes place.

Instead of beginning from a random coil or extended state, folding is “kick started” by native-like topology and intermolecular interactions (Dobson 1992, Blanco *et al.* 1998, Baldwin 2002). For most proteins, polypeptide synthesis requires between 20 seconds and several minutes (Fedorov and Baldwin 1998). As a consequence, compact intermediates are formed in the process of synthesis. Binding of cofactors and ligands, for example oligosaccharides, often stabilizes protein structure and can affect the folding pathways (Kim *et al.* 1991, Chen *et al.* 1995). In the later stages of polypeptide folding, disulfide bonds are formed.

The hydrophobic core of the protein molecule is tightly packed in solution and amino acid changes caused by mutations must be compensated by surrounding neighbors (Chothia 1975). The last stage of protein folding involves the ordering of amino acid side chains into a well-defined and closely packed configuration to maximize noncovalent interactions (Richards 1974). At this point the overall fold has already been determined, dictated largely by hydrophobic interactions in the early stages of folding, and the side chains are close to their final positions (Kuwajima 1996). Overall, protein folding is still not well understood, although some basic principles have been described (Daggett and Fersht 2003).

2.11. Chaperones

Molecular chaperones are ubiquitous components of cells that assist the folding of many proteins. Chaperones decrease protein aggregation by binding aggregation-prone intermediates, unfolding nonproductive folding intermediates, and assisting the polypeptide

in overcoming energy barriers in the folding reaction (Buchner 1996, Fedorov and Baldwin 1997). The majority of the currently known chaperones belong to three highly conserved heat-shock protein families (hsp70, hsp90, and hsp60), whose members are widely distributed from prokaryotes to plants and mammals (Gething and Sambrook 1992). Chaperones are found in the cytosol, ER, mitochondria, and nucleus of eukaryotic cells. These proteins are present in great abundance under normal growth conditions, but can be further induced by heat shock or other forms of stress. Members of the hsp70 family are associated with the nascent polypeptides, recognizing abnormally exposed patches of hydrophobic amino acids on the surface of target proteins (Beckmann *et al.* 1990). BiP, a special hsp70 protein, helps to fold proteins in the ER by stabilizing prefolded structures and aiding in completion of translocation (Vogel *et al.* 1990, Blount and Merlie 1991). BiP associates transiently with nascent wild-type (WT) proteins and more permanently with misfolded or unassembled proteins (Hurtley *et al.* 1989). The bacterial GroEL/GroES system is a well-characterized example of hsp60-aided protein folding (Grallert and Buchner 2001).

Some other chaperones have been implicated in binding nascent chains of particular groups of proteins. Calnexin, a chaperone in the ER membrane, binds transiently to some glycoproteins and is required for their correct folding and assembly (Chen *et al.* 1995). Calreticulin, a soluble protein of the ER lumen, and calnexin bind to sugar moieties on nascent glycoproteins through their globular lectin domain (Hammond and Helenius 1995). Monoglycosylated forms of target proteins bind to calnexin and/or calreticulin, until released in deglycosylated form by glucosidase II, that acts specifically on folded glycoproteins (Hebert *et al.* 1995, Ellgaard and Helenius 2003). UDP-glucose:glycoprotein glycosyltransferase reglycosylates unfolded proteins so that they can reassociate with calnexin/calreticulin. A simple model suggests that proteins bound to chaperones are safe from proteolysis and that the rounds of binding and release from chaperones assist protein maturation (Farr *et al.* 1997). Folded proteins in the ER are released to traverse the secretory pathway, but retained unfolded proteins are captured by the degradative machinery. In the disposal pathway, mannosidase I slowly acts on misfolded proteins and directs them for degradation (Jakob *et al.* 1998, Liu *et al.* 1999). Subsequent interaction with ER-degradation-enhancing 1,2-mannosidase-like protein probably diverts the glycoprotein from the calnexin/calreticulin cycle and promotes its degradation (Hosokawa *et al.* 2001). The ER quality control system is important for the fidelity of cellular functions, as quite a large fraction of the ER-synthesized proteins fail to fold and mature properly. Additionally, an incomplete quality control system underlies several genetic disorders, since proteins with amino acid substitutions often fail to fold correctly, leading to ER retention and degradation of misfolded proteins (Aridor and Hannan 2000, 2002).

Protein disulfide isomerase (PDI), a folding catalyst, facilitates formation of the correct set of disulfide bonds by promoting rapid reshuffling of incorrect disulfide pairings. PDI can catalyze thiol/disulfide interchange reactions and promotes protein disulfide formation, isomerization, or reduction of nascent polypeptides *in vivo* (Roth and Pierce 1987). Peptidyl-prolyl isomerases (PPIases), residing in the cytosol and ER of eukaryotes, form transient complexes with nascent polypeptides and accelerate the folding of proline-containing polypeptides (Klappa *et al.* 1995).

2.2. Protein modification

Common post-translational modifications include proteolytic cleavage, glycosylation, acylation and fatty acid modification, phosphorylation, methylation, sulfation, disulfide bond formation, deamidation, ubiquitination, and acetylation (Mann and Jensen 2003). Glycosylation is an important and very common post-translational modification in eukaryotic proteins. It plays a key role for many proteins during folding and stabilizes proteins (Kaushal *et al.* 1994, Busca *et al.* 1995). In addition, glycosylation has a role in protein targeting, cell-cell recognition and signaling, and it serves other regulatory functions. The glycosyl transferase pathway in the Golgi synthesizes oligosaccharides O-linked to serine, threonine, or hydroxylysine. The predominant type of glycosylation in mammals is N-linked glycosylation, where specific asparagine residues of the polypeptide within the consensus sequence, Asn-X-Ser/Thr (in which X can be any residue except proline), are attached to an oligosaccharide moiety (Marshall 1974). N-linked oligosaccharides are attached to the amide group of asparagine residues in the lumen of the ER by a dolichol-linked pathway. The oligosaccharide moiety is then trimmed in the ER and further modifications and additions occur in the Golgi, depending on the protein (Kornfeld and Kornfeld 1985, Aronson and Kuranda 1989). The N-linked oligosaccharide core consists of two N-acetylglucosamine (GlcNAc) and three mannose residues, shown in Figure 1. In the high mannose type of glycoconjugates the terminal region of the oligosaccharide chain initially consists of six additional mannose residues, while in the complex type oligosaccharides the terminal regions contain GlcNac, galactose, sialic acid, and in some cases fucose residues (Kornfeld and Kornfeld 1976). Hybrid type oligosaccharides are a mixture of these types. Defects of glycosylation may result in disease. Examples include a lethal form of Marfan syndrome resulting from a point mutation which creates an extra consensus sequence for N-glycosylation that appears to be utilized in the patient's cells (Lönnqvist *et al.* 1996). Congenital disorders of glycosylation lead to the development of severe multi-system disorders. Congenital disorders of N-glycosylation are divided into two types: Type 1, where the deficiencies are in the assembly of the core oligosaccharide, and type 2, where the defects are in the modification of the core oligosaccharide (Aridor and Hannan 2002). Defects in O-glycosylation cause, for example, the Ehlers-Danlos syndrome (Quentin *et al.* 1990).

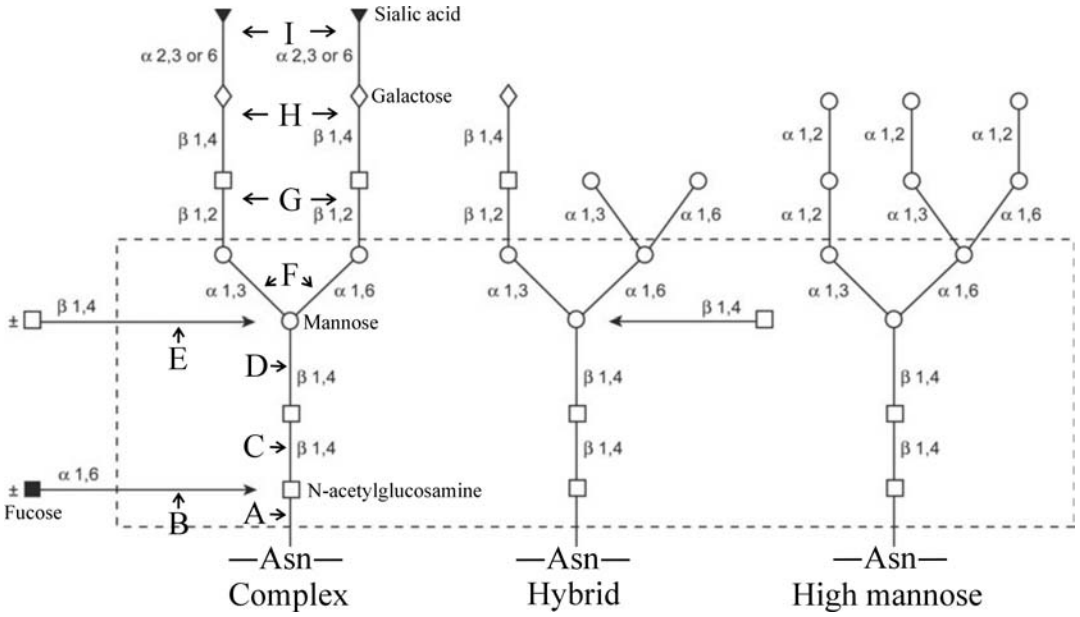

Figure 1. The structure of complex, hybrid type, and high mannose oligosaccharide chains. The lysosomal hydrolytic glycosidases involved in their degradation are: A) aspartylglucosaminidase, B) α-fucosidase, C) endo-β-N-acetyl-glucosaminidase or β-hexosaminidase A or B, D) β-mannosidase, E) β-hexosaminidase A or B, F) α-mannosidase, G) β-hexosaminidase A or B, H) β-galactosidase, and I) α-neuraminidase. The common pentasaccharide core is indicated within a box. Modified with permission from Scriver, C.R, Beaudet, A.L., Sly, W.S., and Valle, D. (Eds.) (2001) The metabolic and molecular basis of inherited disease, pp. 3510, and King, M.W. 2003 (WWW-Page, modified 12.2.2003) The Medical Biochemistry Page, http://www.indstate.edu/thcme/mwking/protein-modifications.html.

2.3. Degradation of misfolded proteins

The cellular folding machinery competes with a different system that transfers proteins containing exposed patches of hydrophobic amino acids to proteasomes for complete destruction. Protein degradation represents the last chance for a cell to suppress the potentially toxic effects of aberrant proteins. It has been suggested that cells stressed by heat shock might die not because of the loss of protein activity due to their denaturation, but because of the toxicity of the denatured proteins (Friant *et al.* 2003). For example, it is essential that non-functional or partially functional ion channels and receptors do not reach the plasma membrane, where their presence could be toxic (Zerangue *et al.* 1999). The cytosolic ubiquitin/proteasome pathway helps clear misfolded proteins (Finley and Chau 1991). Ubiquitin, a small protein, is covalently conjugated onto a lysine residue of a target protein by ubiquitin-conjugating enzyme. Subsequently, more ubiquitin is added to form a polyubiquitin chain. Polyubiquitinated proteins are recognized and rapidly degraded by the

proteasomes (Hiller *et al.* 1996), which are located in the cytosol and the nucleus. Monoubiquitylation seems to function as a reversible non-proteolytic modification that controls protein function in endocytic trafficking (Hicke 2001). Persistently misfolded glycoproteins of the ER lumen undergo ER-associated degradation, where misfolded proteins are retro-translocated to the cytosol and degraded by the proteasomes (McCracken and Brodsky 1996, Kopito 1997). Mutant presenilin-1 protein is known to cause familial Alzheimer's disease. It downregulates the signalling pathway, the unfolded-protein response, that adjusts levels of molecular chaperones in response to stress (Katayama *et al.* 1999). The ER quality control system seems unable to deal with heavily aggregated proteins (Braakman *et al.* 1992, Rivera *et al.* 2000). Aggregates in the ER may compromise cell function, but in many cases may not cause immediate cell death - so they may be similar to inclusion bodies in the cytosol, which are also relatively benign.

3. Intracellular transport of soluble lysosomal proteins

Once secretory proteins are correctly folded, they enter the ER exit sites and are packaged into coatomer protein (COP)-II coated vesicles (Barlowe 2002). Transport from the ER to the *cis*-face of the Golgi occurs in COP-II vesicles via the ER-Golgi intermediate compartment (Saraste and Kuismanen 1984) while transport in the opposite direction (retrograde) occurs in COP-I-coated vesicles (Semenza *et al.* 1990). The ER-to-Golgi movement of transport vesicles, when visualized by live-cell microscopy, occurs along microtubules using the dynein/dynactin motor complex (Presley *et al.* 1997). In the *trans*-Golgi network (TGN), lysosomal enzymes are separated from secretory proteins and transported to the endosomes and then to the lysosomes. Clathrin-coated vesicles mediate traffic between the cell surface, TGN, and the endosomes. Clathrin forms the outer surface of these coats, while adaptor protein (AP) complexes bind clathrin, the coat accessory factors and the cytosolic tails of the transmembrane proteins. The clathrin coat is shed rapidly after vesicle formation. The lysosomal proteins are phosphorylated in the *cis*-Golgi by UDP-N-acetylglucosamine-lysosomal enzyme-N-acetylglucosamine-phosphotransferase, which attaches an N-acetylglucosamine-phosphate group to mannose residues (Hasilik *et al.* 1981, Reitman and Kornfeld 1981, Bao *et al.* 1996a, Bao *et al.* 1996b). Mucolipidosis IIIC has been reported to be caused by mutations in the phosphotransferase γ-subunit gene (Raas-Rothschild *et al.* 2000). The oligosaccharides are further modified in the Golgi, where N-acetylglucosamine-1-phosphodiester-α-N-acetylglucosaminidase removes the GlcNAc group and the lysosomal targeting signal, mannose 6-phosphate, is exposed (Waheed *et al.* 1981). Soluble lysosomal enzymes are sorted into lysosomes by mannose 6-phosphate receptors (MPR) in the *trans*-Golgi (Figure 2).

The transmembrane receptors are themselves directed through intracellular transport pathways by sorting determinants in their cytoplasmic domains. The AP-1 complex mediates MPR sorting at TGN, while AP-3 acts in MPR retrieval from endosomes (Pearse and Robinson 1990, Dell'Angelica *et al.* 1998). AP-1 may also function in retrieval from endosomes. The AP-2 complex functions in endocytic clathrin coated vesicle formation (Robinson 1994, Le Borgne and Hoflack 1998). Two distinct but related MPRs with complementary functions have been described: 300 kDa cation-independent MPR (CI-MPR) (Sahagian *et al.* 1981) and 46 kDa cation-dependent MPR (CD-MPR) (Hoflack and Kornfeld 1985). Both receptors are required for efficient intracellular targeting of lysosomal enzymes and the receptors may interact *in vivo* with distinct but overlapping subsets of phosphorylated ligands (Ludwig *et al.* 1994). MPRs are type-1 integral membrane glycoproteins with mannose 6-phosphate binding sites in the extracytoplasmic region (Dahms and Hancock 2002, Hancock *et al.* 2002). The cytoplasmic tails of both receptors contain numerous sorting signals, some of which are modified by phosphorylation (Rosorius *et al.* 1993) or palmitoylation (Schweizer *et al.* 1996). The CD-MPR undergoes significant structural changes upon ligand binding (Olson *et al.* 2002). The two receptors use a similar carbohydrate-recognition mechanism (Roberts *et al.* 1998, Hancock *et al.* 2002). MPRs loaded with ligands are sorted from the secretory pathway by selective incorporation into tubulo-vesicular structures that emerge from the TGN and ferry cargo to endosomes.

Between early and late endosomes, ligands dissociate from MPRs and are delivered to even more acidic lysosomes, while MPRs are retrieved to the TGN to pick up more cargo (Mellman *et al.* 1986). Extracellular, missorted mannose 6-phosphate-tagged enzymes are recruited by CI-MPR and directed to the endocytic pathway (Helenius *et al.* 1983). CI-MPR also binds and internalizes IGF-II growth factor at the cell surface, thus regulating cell growth (Kiess *et al.* 1988, Wang *et al.* 1994). In the endocytic pathway, incoming material encounters progressively decreasing pH as it moves through the prelysosomal compartments towards the lysosomes. The acidic pH of the lysosomes helps to denature the proteins destined to be degraded. Proteins sequestered within the lysosomes turn over with an apparent half-life of about eight minutes (Mortimore *et al.* 1988). However, lysosomal enzymes have a relatively long half-life in spite of their harsh surroundings. Defects in generation of the lysosomal sorting signal cause diseases. For example, cells affected with the I-cell disease are deficient in GlcNAc-phosphotransferase activity, which results in secretion of soluble lysosomal enzymes from the cells into the extracellular milieu (Reitman *et al.* 1981, Mueller *et al.* 1983).

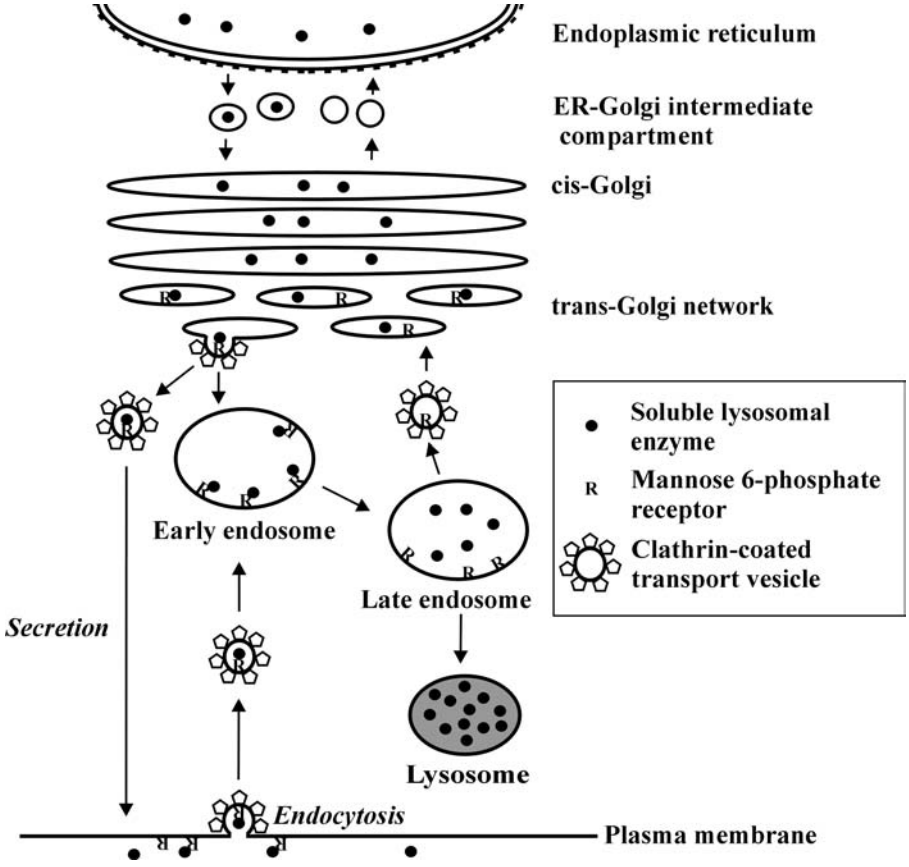

Figure 2. Transport of newly synthesized proteins along the mannose 6-phosphate receptor pathway. Newly synthesized lysosomal enzymes are translocated into the lumen of the ER. The enzymes are transported through the ER-Golgi intermediate compartment to the *cis*-face of the Golgi complex. In the Golgi, N-linked oligosaccharides of soluble lysosomal proteins are labelled with mannose 6-phosphate and the exposed mannose 6-phosphate marker is recognized and bound by the mannose 6-phosphate receptors. The *trans*-Golgi network is the sorting station of the secretory pathway and lysosomal enzymes are delivered to the endosomes. Extracellular lysosomal enzymes are bound to the receptor on the cell surface and endocytosed. In the endosomes, the receptor-ligand complex is broken and the receptors are retrieved and recycled back to the Golgi. Extracellular lysosomal enzymes can be endocytosed from the plasma membrane by mannose 6-phosphate receptors and transported to the lysosomes via endosomes.

4. Aspartylglucosaminidase (AGA)

4.1. Properties and purification of AGA

Aspartylglucosaminidase (AGA, glycosylasparaginase, N^4-(β-*N*-Acetylglucosaminyl)-L-Asparaginase, EC 3.5.1.26) is a lysosomal hydrolase that participates in the breakdown of glycoproteins. Sequential digestion of both the polypeptide backbone and the asparagine-linked oligosaccharide moieties by lysosomal hydrolases results in free amino acid residues and glycosylated asparagines, the latter being substrates of AGA. Specifically, AGA cleaves the amido bond between the N-linked oligosaccharide and the asparagine residue (Makino *et al.* 1966, Tikkanen *et al.* 1996a). AGA only hydrolyses substrates with free α-carboxyl and α-amino groups on L-asparagine but the asparagine-linked carbohydrate chain can be of high mannose, hybrid, or complex type (Makino *et al.* 1968, Tarentino *et al.* 1975, Kaartinen *et al.* 1992). However, fucosylated complex-type N-glycans are not hydrolyzed by AGA (Tarentino *et al.* 1975, Yamashita *et al.* 1979, Noronkoski and Mononen 1997). The hydrolysis reaction of the natural substrate, 2-acetamido-1-β-(L-aspartamido)-1,2-dideoxy-*D*-glucose (AADG), catalyzed by AGA produces aspartic acid and aminoglycan and subsequently the aminoglycan part of the substrate hydrolyzes nonenzymatically to release oligosaccharide and ammonia (Makino *et al.* 1968). The released oligosaccharides are degraded by lysosomal exoglycosidases, which remove sugars from the nonreducing ends of oligosaccharides (Aronson and Kuranda 1989) as illustrated in Figure 1. Specific deficiencies of these enzymes result in glycoprotein or oligosaccharide storage diseases (Cantz and Ulrich-Bott 1990). AGA has other catalytic functions as well. AGA catalyzes the synthesis and degradation of β-aspartyl peptides (Noronkoski *et al.* 1998). In addition, it catalyzes the hydrolysis of L-asparagine, which competitively inhibits the hydrolysis of glycosylated asparagines (Noronkoski *et al.* 1997). Other inhibitors of AGA include L-aspartate and its analogues, transition state mimics, AADG analogues (Risley *et al.* 2001a), and 5-diazo-4-oxo-L-norvaline (Tarentino and Maley 1969, Tarentino *et al.* 1975, Tollersrud *et al.* 1988, Kaartinen *et al.* 1991).

AGA has been partially purified and characterized from the guinea pig serum (Makino *et al.* 1966), the rat liver (Mahadevan and Tappel 1967), the pig serum (Makino *et al.* 1968), the chicken oviduct (Tarentino and Maley 1969), and the pig kidney (Kohno and Yamashina 1972). AGA was purified to homogeneity for the first time in 1989 from human (Baumann *et al.* 1989) and rat (Tollersrud and Aronson 1989) liver. The human AGA enzyme has been purified to homogeneity from the liver (Baumann *et al.* 1989, Rip *et al.* 1992), urine (Kaartinen 1991), and leukocytes (Halila *et al.* 1991, Kaartinen *et al.* 1991, Heiskanen *et al.* 1994, Tollersrud *et al.* 1994a, Tikkanen *et al.* 1996b). The kinetic parameters (Kaartinen *et al.* 1991, Tikkanen *et al.* 1996a, Risley *et al.* 2001b) and specific activity of human AGA

have also been reported (Kaartinen *et al.* 1991). The activity assay for AGA is based on colorimetric measurement of the N-acetylglucosamine liberated from the substrate, AADG (Reissig *et al.* 1955). A sensitive fluorometric assay has also been developed (Mononen *et al.* 1993). The pH optimum of human AGA is between 5.5 and 6 (Baumann *et al.* 1989, Tollersrud and Aronson 1992), while for rat and mouse AGA the optimum is between 7 and 9 (Kaartinen *et al.* 1991, Tollersrud and Aronson 1992). The human enzyme also has high activity between pH values 7 and 9 (Tikkanen *et al.* 1996a). However, at the lysosomal pH, generally below 5, considerable AGA activity also exists (Oinonen *et al.* 1995). The reaction mechanism for AGA-catalyzed amidohydrolysis of glycosylated asparagines is known (Tikkanen *et al.* 1996a) and is well supported by the pH optimum. This has raised suspicions that AGA might have functions outside of lysosomes. For example, involvement in the myelination of the brain neurons has been suggested (Uusitalo *et al.* 1999). The AGA enzyme is stable in the cells, since the half-life of the rat liver AGA is two days (Tollersrud *et al.* 1988). WT AGA is very heat-stable and resistant to denaturation by SDS (Tollersrud and Aronson 1989, Enomaa *et al.* 1992), features utilized in the purification of the active enzyme (Heiskanen *et al.* 1994, Tollersrud *et al.* 1994b, Tikkanen *et al.* 1996b).

4.2. The AGA gene

The human *AGA* gene was physically mapped to the region 4q21-qter of chromosome 4 in 1984 (Aula *et al.* 1984). Later, the gene was linked to a marker at 4q28-31 (Grön *et al.* 1990) and the location was defined by in-situ hybridization methods to 4q32-33 (Morris *et al.* 1992). The 2.1 kb cDNA was cloned and characterized in the early 1990s by two groups (Fisher *et al.* 1990, Ikonen *et al.* 1991b). The first mutation resulting in human aspartylglucosaminuria disease was also identified (Ikonen *et al.* 1991b). Both subunits of the 346 amino acid polypeptide were found to be coded by the single gene of 1041 bp. The AGA cDNA contains a 949 bp 3' untranslated region (UTR) (Ikonen *et al.* 1991b) and human and mouse *AGA* contain approximately 900 bp of a homologous region upstream of the ATG translation initiation codon (Uusitalo *et al.* 1997). The transcription starts 298 upstream of the ATG codon. Two mRNA species (2.1 kb and 1.4 kb) have been identified in human tissues, which could be explained by the use of alternative polyadenylation sites of the 3' UTR (Ikonen *et al.* 1991b). The human AGA enzyme is present in very small quantities in all tissues, as it is encoded by a housekeeping gene (Enomaa *et al.* 1992, Uusitalo *et al.* 1997).

The genomic structure of the 13 kb *AGA* gene was solved in 1991 (Park *et al.* 1991). The gene contains nine exons, several transcriptional elements, and an inhibitory region in the 5' UTR (Uusitalo *et al.* 1997). The *AGA* gene has been found in various organisms from bacteria to mammals: the bacterium *Flavobacterium meningosepticum* (Tarentino and Plummer 1993), the butterfly *Spodoptera frugiperda* (Liu *et al.* 1996), and the mouse

(Tenhunen *et al.* 1995). An evolutionary relationship also exists with a plant asparaginase from the plant *Lupinus arboreus* (Lough *et al.* 1992). In addition, homologs of the human *AGA* gene have been detected in rat (Conchie and Strachan 1969, Tollersrud and Aronson 1989, Liu *et al.* 1996), pig, cow, and chicken (Tollersrud and Aronson 1992, Liu *et al.* 1996). The human genome project (special issues Science 15 February 2001, Nature 16 February 2001, Science 11 April 2003, and Nature 24 April 2003) has provided a sequence of the human genome with 99.9% of the genome covered. In the database, the human *AGA* gene is found in 4q32-33 and the gene spans 10668 bp plus 5' and 3' UTR regions (http://www.ncbi.nlm.nih.gov/, NT 022792). The *AGA* genomic sequence provided by the international consortium contains a few errors; for example, in exon 4 S159 is changed to threonine in the genomic sequence.

4.3. Dimerization and autoproteolysis of AGA precursors

AGA is synthesized as an inactive precursor molecule, which is rapidly cleaved into α- and β- subunits (Ikonen *et al.* 1993). The molecule is simultaneously activated as the critical α-amino group of T206 is exposed. Dimerization of two inactive precursor molecules is suggested to be the event triggering autoproteolytic activation (Riikonen *et al.* 1996). Dimerization takes place either when the polypeptides have reached a conformation that is close to the final folded state or during the early stages of folding. Certain mutations of the AGA polypeptide result in lack of dimerization and subsequent activation (Riikonen *et al.* 1996). The analyzed AGA polypeptides with mutations of the active site are not drastically misfolded as they are transported to the lysosomes (see 4.4) (Peltola *et al.* 1996, Tikkanen *et al.* 1996a). The activation of AGA has been implicated to occur autocatalytically (Oinonen *et al.* 1995, Peltola *et al.* 1996, Tikkanen *et al.* 1996a). Since the activation site of AGA is located at the bottom of the active site funnel, it is unlikely that an external protease could reach the scissile peptide bond (Tikkanen *et al.* 1996a).

4.4. Intracellular processing and transport of AGA

The *AGA* gene encodes an inactive 346-amino-acid precursor polypeptide from which a 23-amino-acid signal sequence is cleaved co-translationally (Figure 3). Two precursor molecules dimerize and the precursor molecules are normally cleaved into the N-terminal α- (27 kDa) and C-terminal β- (17 kDa) subunits (Ikonen *et al.* 1993, Riikonen *et al.* 1995). The molecule also obtains its full enzymatic activity. The AGA polypeptides migrate abnormally on SDS-PAGE due to lack of sufficient binding to SDS (Tollersrud and Aronson 1989, Tollersrud *et al.* 1994a), which has historically created problems regarding the determination of the precise molecular mass. A high-mannose type of oligosaccharide chain is N–linked to each AGA subunit at residues N38 and N308 (Tollersrud and Aronson 1989, Tikkanen *et al.* 1995). The active tetrameric molecule is transported to the

lysosomes, using the mannose 6-phosphate pathway (Tikkanen *et al.* 1995, Park *et al.* 1996b, Tikkanen *et al.* 1997). The same pathway is also utilized by neuronal cells (Kyttälä *et al.* 1998). The C-terminal parts of the human α- and β-subunits are trimmed in the lysosomes by lysosomal proteases (Peltola *et al.* 1996). This processing pattern of AGA can be easily used to monitor the intracellular transport of AGA polypeptides. Some disease-causing mutations of AGA cause abnormal processing of the precursor, where activation does not occur but the molecule is transported to the lysosomes and cleaved (Peltola *et al.* 1996, Tikkanen *et al.* 1996a). Even in these mutations, the AGA molecule is folded sufficiently correctly to be transported to the lysosomes, since the three-dimensional structure determines whether the phosphorylation of AGA oligosaccharides occurs (Tikkanen *et al.* 1997). UDP-N-acetylglucosamine phosphotransferase recognizes lysosomal proteins on the basis of a surface structural determinant. On the surface of the human AGA molecule, the determinant is formed by three co-operating and separate areas containing lysine residues and oligosaccharide chains (Tikkanen *et al.* 1997). The interaction with phosphotransferase is co-operative and all recognition sites have to be in correct spatial position; thus, only correctly folded AGA molecules are phosphorylated and transported to the lysosomes. AGA, like other lysosomal enzymes, is also secreted and subsequently endocytosed by adjacent cells (Enomaa *et al.* 1995). AGA molecules without oligosaccharides are active but are secreted (Tikkanen *et al.* 1995).

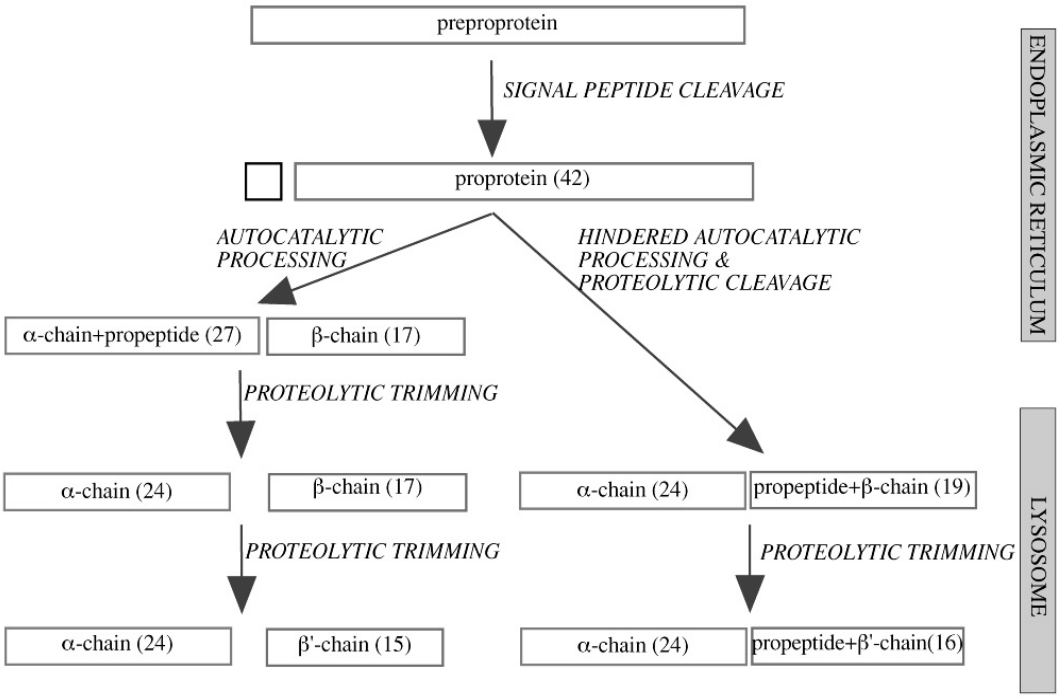

Figure 3. The processing of human AGA. The signal peptide is cleaved co-translationally in the ER from the AGA preproprotein, resulting in a monomeric precursor polypeptide (proprotein). The normal autocatalytic processing pathway is represented on the left, and the abnormal nonactivating pathway caused by processing failure on the right. The apparent molecular weight determined from SDS-PAGE is shown in parentheses. Two precursor molecules dimerize in the ER prior to the activation cleavage, which produces the two αβ-dimers of mature AGA. Copyright (1998) American Society for Biochemistry and Molecular Biology, USA.

4.5. Structure and reaction mechanism of human AGA

The three-dimensional crystal structure of native human AGA was determined at 2.0 Å resolution by Oinonen *et al*. (1995). The active AGA molecule is an $(\alpha\beta)_2$ tetramer with 50*50*70 Å dimensions (1apy, Protein Data Bank, http://www.rcsb.org/pdb/). The molecule is comprised of two almost identical heterodimeric αβ- domains packed together (Figure 4). The core structure of the domains contains an αββα-sandwich, where antiparallel β-strands are packed between α-helical layers. The domains contain four intramolecular disulfide bridges (C64-C69; C163-C179; C286-C306; C317-C345) and are noncovalently bound to each other. The interface between the domains is relatively flat and contains over 200 unique contacts. The upper part of the AGA structure forms a plateau which consists of the C-terminal loops of both α-subunits, the AGU_{Fin}-loops. This loop contains two α-

helices and a disulfide bridge in addition to irregularly folded structures. The heterotetrameric molecule of AGA contains two active sites, one in each domain. The refined crystal structure of AGA in a complex with the reaction product, aspartate, was solved at 2.3 Å resolution (Oinonen *et al.* 1995). The active site is funnel-shaped; the asparagine part of the substrate is bound at the bottom and the oligosaccharide part is bound in the upper and wider part of the funnel, which explains the structural requirements of the substrate. The active site of AGA contains residues from both subunits and the residues are located in loops connecting the different α- and β-layers. T206 functions as the catalytic nucleophile in the substrate hydrolysis reaction and its properties are influenced by hydrogen bonds from S72 and T224. T257 and G258 stabilize the negatively charged reaction intermediate with hydrogen bonds (the oxyanion hole), while the substrate asparagine is bound by hydrogen bonds from T206, R234, D237, T257, and G258 (Oinonen *et al.* 1995). The catalytic site of AGA has no typical serine protease-like catalytic triad (Tikkanen *et al.* 1996a). In the serine proteases, the catalytic triad is formed by Ser (reacting nucleophile) – His (base) – Asp (oxyanion hole: charge relay and stabilization) (Matthews *et al.* 1967, Warshel *et al.* 1989). The key catalytic residue of AGA, T206, is located at the N-terminus of the β-subunit and has a dual function. The prerequisite for the reaction is the uncharged NH_2 state of the α-amino group of T206, which is consistent with the measured pH optimum of AGA (pH 7-9). The reaction is shown in Figure 5. The reaction mechanisms of serine proteases and of AGA have similar features and both serine proteases and AGA use specific and non-specific substrate binding. In addition, the acylation step of the reaction has been proposed to be rate-limiting in AGA catalyzed substrate hydrolysis as well as in serine proteases (Du and Risley 2003).

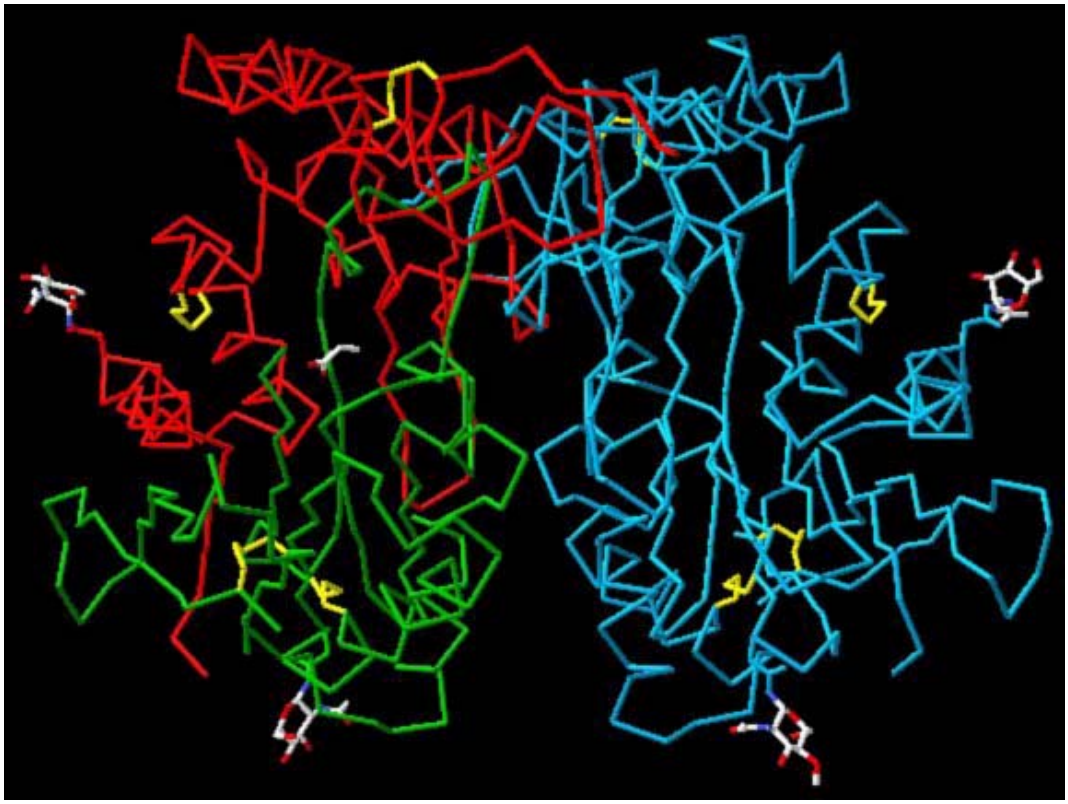

Figure 4. Overall quaternary structure of WT human AGA. One αβ-heterodimer is red (α-subunit) and green (β-subunit), while the other dimer is blue. The disulfide bridges are shown in yellow. The reaction product, aspartate, is shown in one active site. The glycosylation sites, one in each subunit, are shown with one GlcNAc residue. Modified from Oinonen *et al.* 1995.

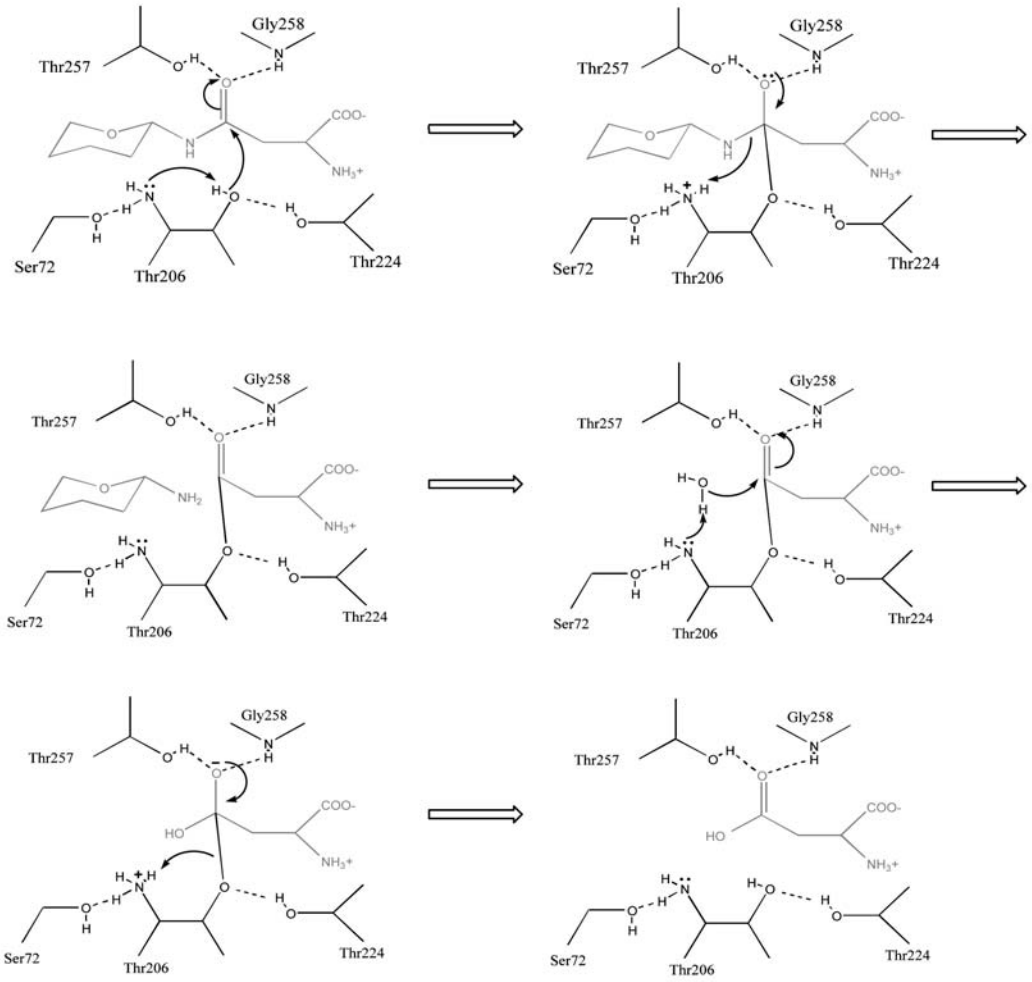

Figure 5. The catalytic reaction mechanism of human AGA. The α-amino group of T206 acts as the base and increases the nucleophilicity of the side chain hydroxyl group of T206, which forms a covalent bond with the carbonyl carbon of the substrate in the acylation step. S72 and T224 fine tune the properties of T206. The negatively charged enzyme-substrate transition state is stabilized by an oxyanion hole, which is formed by hydrogen bonds from T257 and G258. In the deacylation step, a water molecule hydrolyzes the complex, the carbohydrate and aspartic acid are released and the enzyme is regenerated. Modified from Tikkanen *et al*. 1996a.

4.6. The prokaryotic AGA enzyme

The prokaryotic AGA enzyme was purified in 1993 from *Flavobacterium meningosepticum* (Tarentino and Plummer 1993) and the gene was cloned in 1995 (Tarentino *et al.* 1995). The biochemical properties of the bacterial enzyme are similar to the human counterpart. The bacterial AGA enzyme is also inhibited by 5-diazo-4-oxo-L-norvaline. However, the

specific activity is higher than in most mammalian AGAs (Tarentino and Plummer 1993). The periplasmic 295-amino acid bacterial AGA also contains a leader sequence and is synthesized as an inactive precursor which is activated autocatalytically (Tarentino *et al.* 1995, Guan *et al.* 1998, Xu *et al.* 1999). The amino acid sequence is 35% identical with human AGA and the catalytic residues are conserved (Liu *et al.* 1998). The antibodies against the native rat liver AGA cross-reacted with the bacterial AGA (Tarentino and Plummer 1993). The crystal structure of the endogenous WT bacterial enzyme has been published (2gaw, 1ayy) (Guo *et al.* 1998, Xuan *et al.* 1998), and later the precursor form of the enzyme was purified (Cui *et al.* 1999) and the structure solved by crystallography (9gaa, 9gac, 9gaf) (Xu *et al.* 1999). Two other precursor structures were published recently (1p4k, 1p4v) (Qian *et al.* 2003). The folding of the bacterial glycosylasparaginase is very similar to that of the human AGA, although the sequence identity is quite low and disulfide bridges are not present in the bacterial enzyme (Xuan *et al.* 1998). In contrast to the human AGA, no trimming of the C-terminal parts of the subunits occurs in the bacterial enzyme and it is nonglycosylated. The structure reveals that the active site funnel is wider in the bacterial than in the human AGA enzyme, although the active site residues are conserved. No extensive comparisons of the substrate requirements have been reported. The similarity of the native enzyme and the precursor molecule provided direct evidence of the autocatalytic activation, since it would be very difficult for any protease to reach the scissile peptide bond at the bottom of the active site funnel. The autocatalytic reaction mechanism of bacterial AGA has also been presented (Xu *et al.* 1999, Qian *et al.* 2003) and the dimerization of precursor polypeptides has been analyzed (Wang and Guo 2003). The lower stability of the bacterial αββα-tetramer (Xuan *et al.* 1998) may result from absence of the C-terminal random coil of the human AGA, which stabilizes the molecule (Riikonen *et al.* 1995).

5. Aspartylglucosaminuria (AGU)

5.1. Clinical features and diagnosis

The Finnish disease heritage (Perheentupa 1972), a group of rare hereditary diseases that are over-represented in Finland, is a consequence of the population history of Finland (Norio *et al.* 1973, Norio 2003a, 2003b). The heritage comprises at least 36 disorders, most of which are monogenic and autosomally recessive. The second most common of these diseases is aspartylglucosaminuria (AGU, OMIM 208400), a lysosomal storage disease, which is caused by deficient activity of AGA (Pollitt *et al.* 1968, Palo and Mattsson 1970, Autio 1972, Aula *et al.* 2001). The causative mutations were characterized after isolation of the cDNA (Ikonen *et al.* 1991b). The lack of AGA activity causes accumulation of glycoasparagine (mainly AADG) in the lysosomes of all cell types (Haltia *et al.* 1975), resulting in a rather uniform phenotype with progressive mental retardation, coarse facial

features, and skeletal and connective tissue abnormalities (Autio *et al.* 1973). Abnormal amounts of glycoasparagines are also excreted in the urine and can be detected by thin-layer chromatography or liquid chromatography (Pollitt *et al.* 1968, Sugahara *et al.* 1976, Kaartinen and Mononen 1989). The accumulation of a variety of glycoasparagines could result from defective α(1-6)-mannosidase activity. This enzyme only hydrolyzes substrates already hydrolyzed by AGA (Haeuw *et al.* 1994). Virtually all cell types exhibit deficient AGA activity (Palo *et al.* 1972, Maury 1980) but the central nervous system cells are most severely affected. In the brain, AGU primarily affects the grey-matter but delayed myelination in white-matter occurs (Autti *et al.* 1997). The developmental lifespan of the patients is divided into three phases. Firstly, in childhood the patients steadily learn new things. Secondly, in puberty learning is impaired and the patients may lose some skills and abilities. Thirdly, rapid mental and physical retardation occurs after the age of 25-28 years (Arvio and Arvio 2002). Epileptic seizures are common during the later course of the disease. The mean age at diagnosis is 5.5 years in Finland (Arvio *et al.* 1993a), the skills of AGU patients peak at 13-16 years, and the life expectancy is 40 and 35 years in females and males, respectively (Arvio *et al.* 1993b). The most common Finnish mutation, AGU_{Fin} major, manifested itself some 2000-2500 years ago (75-100 generations) and occurs throughout the country (Varilo *et al.* 1996). The linkage disequilibrium in disease alleles can be observed over an interval of 2 centiMorgans, reflecting the relatively old age of the mutation. The disease incidence in Finland is 1:16000 (Peltonen *et al.* 1999), and the carrier frequency varies from 1:132 in Southern Bothnia to 1:63 in North Karelia (Pastinen *et al.* 2001). The existence of the AGU_{Fin} major founder mutation enables the use of rapid and convenient DNA methods for large-scale carrier screening: solid-phase mini-sequencing (Syvänen *et al.* 1992), PCR (Nilssen *et al.* 1993), and DNA-chips (Pastinen *et al.* 2001). More than 260 AGU patients have been reported in Finland and over 30 elsewhere in the world (Norio 2003c).

5.2. AGU mutations

AGU was first reported in 1968 in two English siblings (Pollitt *et al.* 1968). AGU is the most common disorder of glycoprotein metabolism with defective N-linked oligosaccharide breakdown. Other glycoproteinoses include α-mannosidosis (Nilssen *et al.* 1997), β-mannosidosis (Alkhayat *et al.* 1998), and Schindler disease (Wang *et al.* 1990). The AGU_{Fin} major mutation represents 98% of all Finnish disease alleles and originated from one founder mutation (482G>A) which causes an amino acid replacement (C163S) in the AGA polypeptide (Ikonen *et al.* 1991b). Most of the Finnish patients are homozygotes for the AGU_{Fin} major allele. *In vitro* expression of cDNA carrying the AGU_{Fin} mutation has revealed that lack of correct folding and subsequent activation of the AGU polypeptide cause rapid intracellular degradation of the polypeptide (Ikonen *et al.* 1991c, Riikonen *et al.* 1994). The other Finnish mutation, AGU_{Fin} minor, is a 2 bp deletion which causes a

frameshift and a premature stop codon in addition to reduced transcription levels (Isoniemi *et al.* 1995). Patients with this mutation are always compound heterozygotes with AGU_{Fin} major in the other allele. One founder mutation is responsible for all AGU_{Fin} minor alleles, according to haplotype data (Valkonen *et al.* 1999). AGU disease predisposes patients to chronic arthritis (Arvio *et al.* 1998) and AGU carriers may also be predisposed to arthritis (Arvio *et al.* 2002). A total of 18 sporadic AGU alleles have been reported outside of Finland (Fisher and Aronson 1991, Ikonen *et al.* 1991a, Yoshida *et al.* 1992, Park *et al.* 1993, Peltola *et al.* 1994, Jalanko *et al.* 1995, Park *et al.* 1996a, Peltola *et al.* 1996, Laitinen *et al.* 1997, Coulter-Mackie 1999). All known AGU mutations are summarized in Figure 10. In addition, partial deletion of the 3' UTR causes AGU in one case (Ikonen *et al.* 1992). The effects of AGU mutations at the polypeptide level have been reported for some of the mutations: AGU_{Fin} (Ikonen *et al.* 1991c, Enomaa *et al.* 1992, Riikonen *et al.* 1994), C64X (Peltola *et al.* 1994), S72P (Peltola *et al.* 1996), A101V (Park *et al.* 1993), L263X (Park *et al.* 1996a), S269fsX378 (Fisher and Aronson 1991), and G314fsX378 (Jalanko *et al.* 1995). However, the functional link between known molecular genetic and enzyme defects and changes in neuronal cell pathology has remained largely unknown for lysosomal storage diseases including AGU (Walkley 1998).

5.3. AGU mouse model

The mouse *AGA* gene structure is also known (Tenhunen *et al.* 1995). The 11 kb gene contains 9 exons and the 1.2 kb mRNA contains a 1041 bp cDNA whose sequence identity is 84% compared to human. The sequence identity of the mouse AGA polypeptide (also 346 amino acids) is 84% and the polypeptide is processed into α- and β-subunits similarly to the human counterpart. The active site residues are well conserved. Two mouse models of AGU have been generated (Kaartinen *et al.* 1996, Jalanko *et al.* 1998). The phenotype and histopathology of the AGU mice mimic the human disease characteristics well (Jalanko *et al.* 1998, Kaartinen *et al.* 1998, Tenhunen *et al.* 1998). The latter mouse model has been utilized in bone-marrow transplantation (Laine *et al.* 1999) and gene therapy (Peltola *et al.* 1998, Harkke *et al.* 2003) experiments with promising results.

6. N-terminal nucleophile hydrolases

The core αββα -structure, the N-terminal nucleophile, and the topology of secondary elements characteristic of AGA have been found in other proteins as well (Brannigan *et al.* 1995). The N-terminal nucleophile hydrolase (Ntn-hydrolase) enzyme superfamily consists of AGA (Oinonen *et al.* 1995), the β-subunit of the proteasome (Löwe *et al.* 1995), penicillin G acylase (Duggleby *et al.* 1995), penicillin V acylase (Suresh *et al.* 1999), glutamine phosphoribosylpyrophosphate (PRPP) amidotransferase (Smith *et al.* 1994),

glutaryl 7-aminocephalosporanic acid acylase (Lee *et al.* 2000), cephalosporin acylase (Kim *et al.* 2000), and glucosamine-6-phosphate synthase (Isupov *et al.* 1996). So far, AGA and the proteasome are the only mammalian members of the superfamily which have been characterized. The activation of the enzymes occurs autocatalytically from an inactive polypeptide and reveals the critical α-amino group of the catalytic residue, which is Thr, Ser, or Cys. The side chain of the catalytic residue, the nucleophile, catalyzes the hydrolysis of a substrate amido bond and each enzyme has a different substrate specificity. In the case of AGA, penicillin G acylase, cephalosporin acylase, and glutaryl 7-aminocephalosporanic acid acylase activation, the precursor polypeptide chain is cleaved into two polypeptide subunits of the active protein, whereas in proteasome and glutamine PRPP amidotransferase, the activation results in the removal of a propeptide.

A range of autocatalytic events proceed via protein splicing (Perler *et al.* 1994), where the initial steps are thought to involve an N->O acyl rearrangement to generate a branched intermediate (Shao *et al.* 1996). In Ntn-hydrolases, the hydrolysis of the peptide bond preceding the N-terminal active site residue by an autocatalytic rearrangement is similar to the initial step in the protein splicing pathway. The N->O acyl rearrangement reaction is strongly pH-dependent, favoring the ester under acidic conditions and the amide under alkaline conditions. In Ntn-hydrolases, it is generally believed that the catalytic nucleophile also acts as the autocatalytic nucleophile (Schmidtke *et al.* 1996, Lee and Park 1998, Li *et al.* 1999, Hewitt *et al.* 2000, Kim *et al.* 2002). Other active site residues may also have roles in both the catalytic and autocatalytic reactions.

6.1. Proteasome β-subunit (PRO)

The proteasome is the central enzyme of protein degradation in the cytosol and the nucleus (EC 3.4.25.1). It is composed of regulatory α-subunits and proteolytic β-subunits (PRO). The subunits are arranged as 4 rings: the outer rings contain 7 α- and the inner rings 7 β-subunits. The 7 PRO polypeptides are all different and only 3 of them are active with different substrate specificities (Dick *et al.* 1998). The archaebacterial (1pma) (Löwe *et al.* 1995) crystal structure of the active complex has been solved, whereas the yeast complex (1ryp) (Groll *et al.* 1997) contains mutagenized active site residues (Chains H and V: T1A; Chains L and Z: K33R). The 28-member complex is barrel-shaped and the active sites are on the inner surface. The precursor of PRO is activated by removal of the propeptide, which reveals the critical α-amino group of the catalytic residue, T1. For the autocatalytic reaction, the N-terminal amino group is not available as a proton acceptor. K33 has been considered to act as the general base (Schmidtke *et al.* 1996), although it is stabilized in its protonated form by three well-defined hydrogen bonds and a salt link to D17 in yeast (Ditzel *et al.* 1998). Instead, a catalytic water molecule, seen in the high-resolution structure of three active subunits of the yeast proteasome (Groll *et al.* 1997), is

ideally positioned to act as the general base and promote the abstraction of the proton from the T1 hydroxyl group. This enables nucleophilic attack on the carbonyl carbon of the preceding peptide bond. The resulting intermediate decays to the ester, which is subsequently hydrolyzed to release the propeptide (Ditzel *et al.* 1998). The propeptide adopts a strained turn, the γ-turn, (Milner-White 1990) between L-2 and T1. S129 stabilizes the transition state during the activation, while G47 performs this role in enzyme catalysis (Ditzel *et al.* 1998).

6.2. Penicillin G acylase (PGA)

The *in vivo* role of the bacterial penicillin G acylase (PGA) (EC 3.5.1.11) is not yet clear, but the enzyme is used on an industrial scale to produce antibiotics (Bruggink *et al.* 1998). The native penicillin G acylase crystal structure from *Eschericia coli* has been solved (1pnk) (Duggleby *et al.* 1995), and further re-refined to 1.3 Å (1gk9) (McVey *et al.* 2001). The native enzyme is an αβ-dimer. Activation of penicillin G acylase occurs autocatalytically (Choi *et al.* 1992, Kasche *et al.* 1999). The precursor (1e3a) and mature forms of PGA have very similar structures (Hewitt *et al.* 2000). However, substantial structural differences exist in the C-terminal part of the α-chain. In the precursor, the linker peptide is cleaved autocatalytically between Y260 and P261 by O^{γ} of T262 according to dynamic simulation. The catalytic residue, S264, has two well-defined conformations that can be correlated to the presence or absence of the linker peptide. The same residues that maintain the tetrahedral intermediate in the active enzyme are also suggested to be responsible for the stabilization of the autocatalytically generated intermediate of precursor activation. However, the B-factors of the residues at the processing site are high and the stereochemistry at this site may differ from that of the mature enzyme (Hewitt *et al.* 2000). The peptide bond between residues G263 and S264 of the mutant precursor shows non-ideal stereochemistry and is consistent with partial cleavage, disorder or distortion of this bond. The three residues preceding the nucleophile are important for efficient processing of the precursor (Choi *et al.* 1992). R526 is crucial for efficient processing and stabilization of the oxyanion hole (Alkema *et al.* 2002), confirming that an intact oxyanion hole is necessary for efficient processing of the precursor enzyme (Alkema *et al.* 2000).

Penicillin V acylase (PVA) and penicillin G acylase (PGA) have similar substrates, but their molecular properties are different. PVA from *Bacillus sphaericus* (2pva, 3pva) is tetrameric, while PGA from *E. coli* is a heterodimer. The enzymes have no sequence similarity. They share the αββα-structural core but are otherwise unrelated in primary sequence, including the active site residue (Duggleby *et al.* 1995, Suresh *et al.* 1999).

6.3. Glutamine amidotransferase (GAT)

Glutamine 5-phosphoribosyl-1-pyrophosphate amidotransferase (EC 2.4.2.14) catalyzes the initial reaction in de novo purine nucleotide biosynthesis and is the key regulatory enzyme in the pathway. The tetrameric native enzyme from *Bacillus subtilis* contains two structural domains: an N-terminal glutaminase domain (GAT) and a C-terminal phosphoribosylamine-synthesizing domain (1gph) (Smith *et al.* 1994). The GAT domain belongs to the Ntn superfamily and is activated by propeptide removal. The functional residues of the *E. coli* GAT include the nucleophile (C1), the oxyanion hole (N101 and G102), substrate binding (R73 and D127), and interdomain coupling (Y74) (Kim *et al.* 1996). Mutational analysis of *B. subtilis* GAT (1ecf) (Muchmore *et al.* 1998) showed that the propeptide was not processed after replacement of the nucleophile or two residues of the oxyanion hole (Li *et al.* 1999). The results suggest that the oxyanion hole which participates in catalysis also has a role in propeptide processing, possibly by stabilizing the tetrahedral intermediate resulting from a nucleophilic attack of C1 thiolate on the peptide bond between E-1 and C1. A general acid or base is also needed for propeptide cleavage. The α-amino group of C1 is in peptide linkage and is therefore unavailable to assume this function. For this purpose, mutagenesis of conserved histidine residues was performed and the results revealed that H70 was necessary to maintain propeptide processing (Li *et al.* 1999). However, no direct role for H70 could be assigned. In addition, mutant polypeptides carrying several replacements of E-2 and a double mutant of residues E-2/E-1 were partially processed. The interpretation of these results leads to the conclusion that some of the residues needed for enzymatic catalysis are involved in propeptide processing. In the active site of the mature *E. coli* enzyme, the oxyanion hole residues and C1 are hydrogen-bonded to the same water molecule (Muchmore *et al.* 1998).

6.4. Other members of the superfamily

Other Ntn-hydrolases include glutaryl 7-aminocephalosporanic acid acylase (EC 3.5.1.11) (Lee *et al.* 2000) and cephalosporin acylase (3.5.1.11) (Kim *et al.* 2000), which have an N-terminal serine in the β-subunit. The enzymes are functionally related to PGA and the active sites are structurally similar. However, the enzymes have different substrates (Fritz-Wolf *et al.* 2002). Cephalosporin acylase (CA) catalytic amino acids have been characterized (Kim and Kim 2001). The active αβ-heterodimeric molecule (1fm2) (Kim *et al.* 2000) and the precursor molecule from *Pseudomonas diminuta* have very similar structures (1keh) (Kim *et al.* 2002). Both structures contain two water molecules, which assist S1β in carrying out the nucleophilic attack in intramolecular activation of the precursor. The role of the second water molecule has been suggested to be stabilization of the reaction intermediate (oxyanion hole), while the first water molecule acts as a general

acid and base. The first water molecule is bound by four hydrogen bonds in pseudotetrahedral geometry (Kim *et al.* 2002).

The analysis of the activation mechanism of glutaryl 7-aminocephalosporanic acid acylase (GCA) from *Pseudomonas* revealed that S199 is essential for catalysis and intramolecular autocatalysis (Lee and Park 1998). Subsequently, the precursor is cut in another location as well, probably by the activated S199 of the other active site of the $(\alpha\beta)_2$ heterotetrameric molecule. The crystal structures of the native (1or0) and the precursor (1oqz) form have been solved but are not released until in March 2004 (Kim *et al.* 2003). Dimerization of the GCA precursors sequesters a slightly larger surface area than in the WT molecule. The interaction between the precursor monomers is based on van der Waals interactions, hydrogen bonds, and salt bridges formed by conserved residues. Dimerization may be guided by electrostatic repulsion of surface patches which exhibit negative potential on the precursor. The structure of the linker peptide between the α- and β-subunits reveals significant differences between the CA and GCA, although the sequences are identical. The CA spacer peptide shows a bend conformation, whereas GCA contains α-helical and β-strand conformation. The substrate binding cavity is shielded by the spacer peptide. After the autocatalytic activation, the nucleophile, S170, at the N-terminus of the β-subunit is exposed to the solvent and the α-amino and side chain groups are hydrogen-bonded to a water molecule. Another water molecule is also trapped in a hydrophobic environment. A reaction mechanism for the activation of the precursor is also suggested. The trapped water molecules may play important roles in the activation.

6.5. Autoproteolytic activation of other proteins

Intramolecular autoproteolysis has also been suggested by kinetic analysis for the first processing step in the activation of non-Ntn-hydrolases such as calpain (Zimmerman and Schlaepfer 1991), papain (Vernet *et al.* 1991), pepsin (Bustin and Conway-Jacobs 1971), mouse prohormone convertase (Goodman and Gorman 1994), and thrombin proenzymes (Petrovan *et al.* 1998), and the maturation of hedgehog (Lee *et al.* 1994, Hall *et al.* 1997) and viral (McCann *et al.* 1994, Zlotnick *et al.* 1994, Wu *et al.* 1998) proteins. The folding of the bacterial serine protease subtilisin is dependent on its prosegment, which is autocatalytically cleaved to give the mature enzyme (Gallagher *et al.* 1995). The biological role of the prosegment is to provide a folding pathway, acting as an intramolecular folding guide. In addition, the activation of inteins occurs autocatalytically (Perler *et al.* 1997). The ester intermediate captured during the activation of aspartate decarboxylase provides evidence of an N->O acyl shift (Albert *et al.* 1998).

AIMS OF THE PRESENT STUDY

This study was carried out in order to characterize in detail the activation process of the human AGA. This knowledge is essential in order to define the detailed cellular consequences of various AGU mutations at the cellular level. To elucidate the disease-causing mechanism of AGU, the consequences of different AGU mutations were also monitored. Finally, the precursor molecule of AGA was expressed and purified in order to solve its three-dimensional crystal structure.

The specific aims of this study were:

1. To characterize the autocatalytic activation mechanism of AGA by *in vitro* expression studies.

2. To identify the amino acids that contribute to the oligomerization of the precursor molecules in the ER.

3. To characterize the consequences of known AGU mutations by *in vitro* expression studies.

4. To express and purify the AGA precursor molecule for structural determination experiments.

MATERIALS AND METHODS

Material or method	Original publication
Patients	II
Genotyping	II
DNA extraction and sequencing	I, II, III
In vitro site-directed mutagenesis	I, II, III
Mutation analysis	I, II, III
Cell culture	I, II, III
Transfections	I, II, III
Metabolic labelling	I, II, III
Immunoprecipitation	I, II, III
Gel electrophoresis	I, II, III
Endoglycosidase H digestion	I
Gel filtration	I, III
Immunofluorescence microscopy	II, III
Assay for AGA activity	I, II, III
Determination of protein concentration	I, II, III
Kinetic measurements	III

Table 1. Materials and methods used in this study.

7. Stable expression of AGA in CHO cells (I)

Chinese hamster ovary cells (CCL61, ATCC) were maintained in DMEM supplemented with 10% fetal calf serum and antibiotics and co-transfected with mutagenized *AGA* cDNA in SV-poly and pCI-neo expression vector (Promega), using the lipofectamin method as suggested by the manufacturer (Life Technologies). Forty-eight hours post transfection the cells were plated with medium containing 500 µg/ml Geneticin (G418, Gibco). After 14-20 days, colonies resistant to G418 were selected and cloned by limited dilution and the clones with the highest AGA activity were selected for further analyses.

8. Production of the yeast cell line for AGA precursor expression

Production of the cell line was performed with the *Pichia* Expression Kit, as suggested by the manufacturer (Invitrogen). The *AGA* cDNA was mutagenized (D205G+T206A) and cloned into the pPIC3.5K expression vector, which allows multiple tandem integrations of the target gene in the *Pichia* genome using homologous recombination. The target gene is under the control of the alcohol oxidase promoter, which can be induced with methanol. The vector was transformed into TOP10F' bacterial cells and *AGA*-containing clones were identified with PCR. Plasmid DNA was extracted from a clone and the sequence was confirmed. The plasmid DNA was digested separately with SacI and SalI restriction enzymes for insertion at the alcohol oxidase 1 gene and the histidinol dehydrogenase gene, respectively, and purified by phenol extraction. The DNA was then transformed into two His^- (histidine-synthesis-deficient) yeast cell lines (KM71 and GS115), using

electroporation. The cells were plated and His^+ transformants were selected. Next the clones were screened for Mut^+ (methanol utilization fast, GS115) and Mut^S (methanol utilization slow, KM71) transformants. To screen for multiple inserts, the transformants were plated on YPD-G418 plates containing increasing concentrations of the neomysin analog G418, against which the vector confers resistance. All G418 resistant clones were picked and purified and the level of G418 resistance was confirmed. Five clones were incubated on methanol-containing plates to promote AGA precursor expression and the AGA expression level was determined. Two clones were chosen for expression and partial purification on a larger scale. Finally, one clone with a 3.0 mg/ml resistance to G418 was chosen for AGA precursor production (GS115 with multiple copies of the *AGA* gene integrated into the alcohol oxidase 1 gene).

9. Expression, purification, and analyses of AGA precursor

The chosen clone expressed AGA precursor approximately at the level of 14 mg/l and secreted AGA into the culture medium using the signal sequence of AGA for targeting to the secretory pathway. The AGA expression was carried out at +30^{o}C in a shaking incubator at 200 rpm. The yeast was first grown to a suitable density in buffered methanol complex medium over four nights. For protein expression, the medium was changed to buffered minimal methanol medium supplemented with 1 mM EDTA (to inhibit metalloproteases) and regular additions of methanol (to promote expression). In addition, 10 mM glycine was added to stabilize the AGA precursor (Cui *et al.* 1999, Xu *et al.* 1999). The cell culture medium was collected and frozen. Using tangential flow filtration (Mini Ultrasette, Filtron Technology Corporation) the medium was concentrated and the buffer changed to buffer suitable for hydrophobic interaction chromatography. Protease inhibitor cocktail (Sigma) was added and the buffer was exchanged for each of the purification steps, using microconcentrators (Centriprep and Centricon, Amicon). Precursor purification was performed using liquid chromatography on a FPLC system (Merck Hitachi) with Pharmacia columns (Amersham Biosciences) at +20^{o}C with ice-cold buffers to minimize AGA precursor cleavage into abnormal subunits. Hydrophobic interaction chromatography was performed on a 5 ml Octyl Sepharose 4 Fast Flow column. Anion exchange chromatography was performed using two MonoQ HR 5/5 columns connected in series. Gel filtration chromatography was performed using a Superdex 200 HR 10/30 column and finally cation exchange chromatography was performed on two MonoS HR 5/5 columns connected in series. Proteins on SDS-PAGE were visualized using Coomassie Brilliant Blue R250 staining (Amersham Biosciences), Sypro Red fluorescent gel staining (FMC), silver staining (SilverXpress, Invitrogen), western blotting, or enhanced chemiluminescence. The purity and properties of the AGA precursor preparation were analyzed using 2-dimensional SDS-PAGE at pH 4-7 (IGPphor, Amersham Biosciences). A mass spectrometric analysis of the preparation was performed at Viikki Biocenter, using an ESI-quadrupoltime-of-flight instrument (Q-TOF1, Micromass).

RESULTS AND DISCUSSION

Due to extensive characterization of human AGA, it has become one of the best-characterized lysosomal enzymes. The intracellular processing, structure, catalytic mechanism, and transport of AGA have been analyzed in detail. However, only preliminary data of the initial events which occur in the ER, including AGA activation and dimerization of precursor molecules, had been obtained prior to this study. This study also elucidates the consequences of patient mutations in detail.

10. Analysis of conserved residues of the AGA molecule

Based on the information of the 3D-structure of human AGA, conservation of the residues of the active site, lysosomal targeting, and glycosylation of asparaginases and glycoasparaginases among 10 species were analyzed. The important residues, regarding the AGA enzyme activity, were found to be conserved from bacteria to mammals. This alignment revealed only 12% overall sequence identity (33/275 residues) and without the plant asparaginases, the sequence identity rose to 26% (71/278). The human AGA sequence is shown in Figure 6. Based on the alignment, two clusters of amino acids showed high conservation in the three-dimensional structure of the native human AGA. One of these regions was located on the interface between the two αβ-dimers, and another on the active site area, implying functionally essential roles for these areas of the molecule. These regions were selected as targets of *in vitro* mutagenesis followed by functional analysis of the mutated polypeptides. As a general practice, the replacing residues were chosen to be of equal size or smaller than the original residue so that the packing of the molecule should not be disturbed. This practice was not followed with some replacements at the dimer interface, where a bulkier side chain was introduced in order to study whether a residue occupying more space affects the dimerization of precursor polypeptides. Where needed, corresponding bacterial amino acids are shown in parentheses.

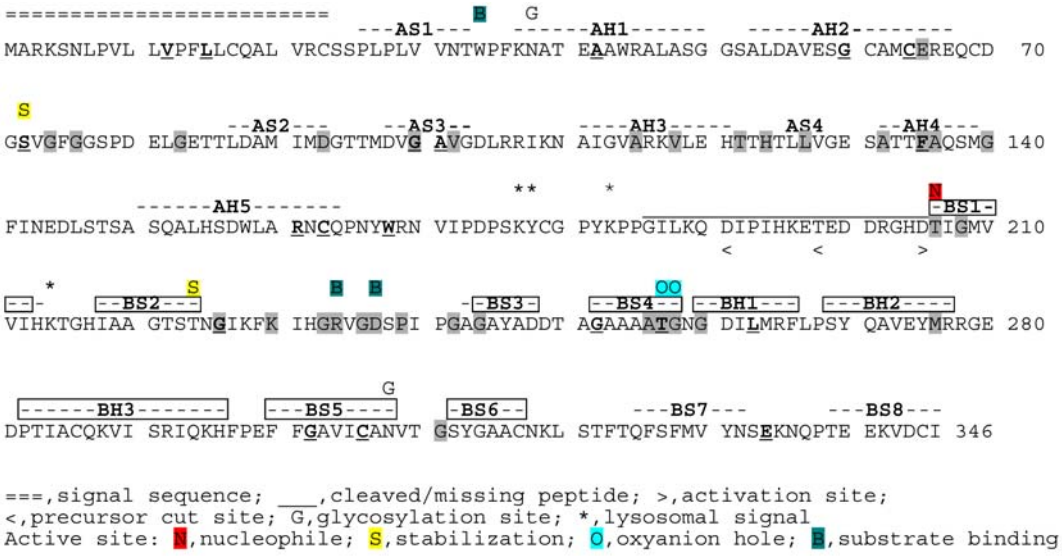

Figure 6. The human AGA amino acid sequence. The residues conserved from plants to mammals are shaded in grey, the locations of patient mutations are bold underlined, and locations of the secondary structure elements are shown above the sequence. The secondary structure elements conserved in Ntn-hydrolases are boxed.

10.1. Dimerization of the precursor molecules (I, II)

Dimerization of AGA precursors occurs post-translationally in the ER and is a prerequisite for AGA activation. The precursor molecules are autocatalytically cleaved into the N-terminal α- (27 kDa) and C-terminal β- (17 kDa) subunits (Ikonen *et al.* 1993, Riikonen *et al.* 1995) thus forming the active tetrameric molecule. The dimer interface of AGA is a relatively wide area containing a number of hydrophobic residues. In the 3D-structure, the interface is quite flat except for one clear protrusion of a loop that penetrates into the other αβ-dimer of the native enzyme. In this loop region, H124 is completely conserved and T125 is conserved in mammals. Two substitutions were found to affect the dimerization of AGA, namely H124R and H124W. Due to a processing failure, these mutants produced only the non-processed precursor polypeptide, which was secreted into the culture medium and thus no enzyme activity beyond the background level was detected. Both of these substitutions affected the dimerization of the precursor molecules and both dimeric and monomeric precursor could be observed after gel filtration. The large side chain of arginine or tryptophan obviously did not fit properly into the narrow gap between the two precursor molecules. When labelled with [^{35}S]Cys and [^{32}P]P_i the oligosaccharides of mutant precursors were found to be normally phosphorylated. This would suggest that these substitutions do not result in severe misfolding of AGA in the ER, since only correct positioning of the three phosphotransferase recognition regions results in phosphorylation of the oligosaccharides of AGA (Tikkanen *et al.* 1997). Correct folding is further supported by the fact that the most common mutation causing aspartylglucosaminuria, AGU_{Fin},

prevents the correct folding of the C-terminal end of the α-subunit (Ikonen *et al.* 1991c, Oinonen *et al.* 1995), and the mutant polypeptides are not phosphorylated *in vitro* and remain monomeric (Peltola *et al.* 1996). In addition, severely misfolded AGA polypeptides can only be immunoprecipitated after the samples have been denatured by boiling and subsequently using antibodies against the denatured subunits (Tikkanen *et al.* 1995). Immunofluorescence analysis of cells expressing H124R and H124W mutants revealed staining of the Golgi and ER, implying that even at steady state, the mutant AGA polypeptides were not found in lysosomes. Western blotting, using antibodies against denatured subunits of AGA, further demonstrated that a small quantity of degradation products existed intracellularly, suggesting instability of the AGA polypeptides carrying these substitutions.

In conclusion, H124 seems to be crucial for the correct packing of the two precursor polypeptides against each other, as suggested by the crystal structure. Substitutions for T125 did not have any effect on the dimerization but partial misfolding and misprocessing was observed. To exclude the effect of over-expression on the results obtained, stable CHO cell lines producing the H124R mutant AGA polypeptides were constructed. The findings in the stable expression system confirmed the results obtained with transient expression in COS-1 cells.

Three patient mutations, G100E, A101V, and F135S, reside on or near the interface between the two αβ-dimers of AGA and were found to affect the dimerization of AGA. The structural environment of F135 is hydrophobic and substitution of F135 for hydrophilic serine eliminates the local stabilizing van der Waals interactions. The delayed activation seen in biochemical experiments probably results from impaired dimerization in the ER, caused by the lack of stabilizing interactions between the two precursor polypeptides. G100E and A101V substitutions disturb the 3D-structure of AGA because of space restrictions and probably also because the local structure around H124 becomes distorted.

Among the patient mutations, P268fsX319, S269fsX274, G314fsX378, and E334X produced more stable polypeptides than C306R or L263X. This indicates that residues D261 to L267 might be important for polypeptide folding and stability. This hypothesis is supported by the AGA structure: these amino acids, like H124, are located between the αβ-dimers of AGA and therefore might stabilize the interaction between two αβ-dimers in the molecule. In addition, Riikonen *et al*. (1996) showed that the large loop structure consisting of the C-terminal part of the α-subunit is crucial for dimerization. Figure 7 shows the most critical regions at the dimer interface.

The dimerization of bacterial AGA has also been analyzed (Wang and Guo 2003). Some conserved residues at the dimer interface, V128, Ile240, E274, and R292 (V105, I186,

E220, and R238), were found to affect the dimerization. These residues are largely buried in the dimer but would be exposed in a monomer and their mutagenesis changes the quaternary structure of the molecule. Previously, Ile240 was found to be essential for human AGA precursor dimerization (Riikonen *et al.* 1996). A two-step dimerization is suggested to trigger bacterial AGA activation (Wang and Guo 2003). In the first step, the folded precursors dimerize through an extensive hydrophobic surface. In the second step, the dimer interface is modified and the dimers interact more strongly, while a conformational strain is created at the activation site. As a result, autoproteolysis is triggered to remove the strain of the polypeptide backbone and generate the N-terminal threonine.

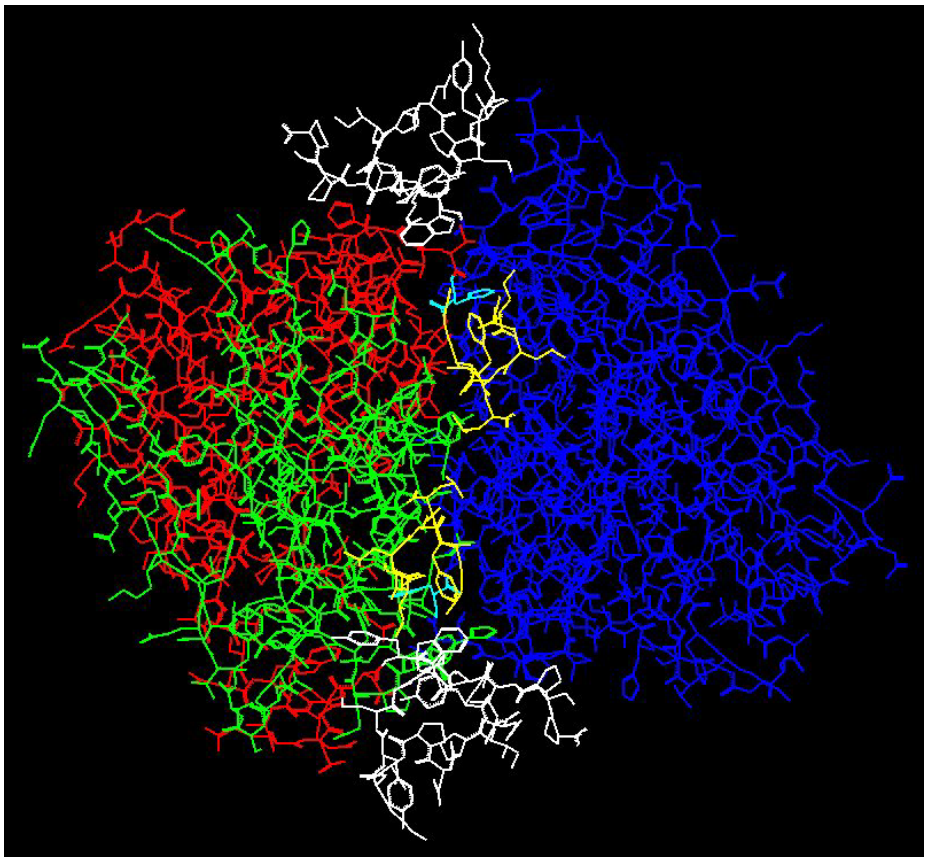

Figure 7. The three-dimensional structure of the native human AGA enzyme. One half of the AGA molecule, comprising one α- and one β-subunit, is red and green and the other is blue. Three regions in both αβ-dimers provide an abundance of contacts between the two αβ-dimers of AGA: Firstly, the loop containing residue H124 (cyan); secondly, the amino acids in the α-helix BH1 and the following loop (yellow); thirdly, the large loop structure at the C-terminal end of the α-subunit (white).

10.2. Autoproteolysis of the precursor molecules (I, III)

Autocatalytic activation of AGA is dependent on correct folding and dimerization of precursor polypeptides. This imposes strict structural requirements for the precursor molecules. Thus even missense mutations of the *AGA* gene often result in severe problems of activation. Prior to this study, activation of the human AGA precursor had been prevented with active site amino acid replacements (Ikonen *et al.* 1993, Riikonen *et al.* 1995, Peltola *et al.* 1996, Tikkanen *et al.* 1996a). The human AGA crystal structure implied that the activation should occur autocatalytically, since the activation site was located in the bottom of the funnel-shaped active site (Oinonen *et al.* 1995). Additionally, some data on endogenous bacterial AGA activation had been obtained (Guan *et al.* 1996). In this study, the residues of the activation site were first replaced, using site-directed *in vitro* mutagenesis (I, 1998). Later, residues surrounding the active site were chosen as targets for mutagenesis (III, 2003). Between these studies, the bacterial AGA was characterized using mutagenesis (Guan *et al.* 1998, Liu *et al.* 1998), and the bacterial AGA WT and precursor crystal structures were published (Guo *et al.* 1998, Xu *et al.* 1999).

10.2.1. Residues participating in autocatalysis (I)

During the autocatalytic activation of AGA, the peptide bond between D205 and T206 is cleaved. At this activation site, three residues were mutagenized: H204 and D205 before the cleavage site and the N-terminal nucleophile T206 immediately after the cleaved bond. Among the mutations at the activation site, only H204G and H204S were partially activated. Other mutations of H204, D205, or T206 totally prevented the activation cleavage. The dimerization of the precursors was not affected by any of the activation site mutations. Apparently, the polypeptides were correctly folded, since they were transported to the lysosomes, where the precursor was processed into α- and pro-β-subunits. Additionally, the cells efficiently secreted precursor polypeptide into the culture medium. Similar processing has been reported with substitutions for amino acids R234 and T257 (Tikkanen *et al.* 1996a). To exclude the effect of over-expression on the results obtained, stable CHO cell lines producing T206C mutant AGA polypeptides were constructed. The findings in the stable expression system confirmed the results obtained by transient expression in COS-1 cells.

The sequence H204-D205-T206 at the proteolytic activation site is conserved in the currently known glycoasparaginases or asparaginases except in plants. The lack of complete conservation would suggest that the side chains of the residues before the cleavage point do not directly participate in the autocatalytic activation in (glyco)asparaginases. Guan *et al.* (1996, 1998) suggested that H204 (H150) would be necessary for initiation of the nucleophilic attack resulting in the activation of bacterial

AGA. In contrast, the results suggest that H204 is not absolutely required for the cleavage reaction, since H204G and H204S mutant precursor polypeptides were activated to some extent. Moreover, the H204S precursor polypeptide has been demonstrated to be cleaved, probably autocatalytically, in the medium (Peltola *et al.* 1996) irrespective of serum or protease inhibitors.

The results show that D205 and T206 are essential for AGA activation and substitutions of these amino acids systematically result in the processing failure described above. In contrast to previous studies (Fisher *et al.* 1993, Guan *et al.* 1996), the results demonstrate that T206 (T152) is required for the actual proteolytic activation of the precursor polypeptide, in addition to its role in the catalytic activity of AGA (Oinonen *et al.* 1995). When T206 of human AGA was replaced by serine, activation cleavage occurred, but the mutant enzyme was inactive (Riikonen *et al.* 1995), indicating that the hydroxyl group of the T206 or S206 side chain is actually required for activation. Only with threonine at the N-terminus is this hydroxyl group in the correct position for enzymatic catalysis in glycoasparagine breakdown.

During activation, the tetrahedral transition state of AGA (Figures 9b and 9e) may be stabilized by one or two hydrogen bonds, the so-called oxyanion hole. In Ntn-hydrolases, several oxyanion hole-forming residues have been proposed (Schmidtke *et al.* 1996, Ditzel *et al.* 1998, Li *et al.* 1999, Xu *et al.* 1999, Hewitt *et al.* 2000, Kim *et al.* 2002, Kim *et al.* 2003). For the activation, an oxyanion hole is required in both the acylation and deacylation steps. In bacterial precursor AGA, T224 (T170) may stabilize the reaction transition state by forming the oxyanion hole during activation (Figures 9b and 9e) (Xu *et al.* 1999). Previous studies of human AGA suggest that a hydroxyl group at T224 is important for autoproteolysis (Tikkanen *et al.* 1996a) and thus T224 may stabilize the reaction transition state by forming the oxyanion hole. However, as the precursor structures of bacterial AGA differ at the scissile D205-T206 peptide bond, it is unclear whether T224 truly stabilizes the reaction intermediate or whether it only stabilizes the active site structure.

10.2.2. Human versus bacterial AGA precursor

The 3D-crystal structure of the human AGA precursor has not been solved so far. However, the existing bacterial AGA precursor structures from the endogenous *F. meningosepticum* enzyme (9gaa, 9gac, 9gaf) (Xu *et al.* 1999), (1p4k, 1p4v) (Qian *et al.* 2003) have provided some clues about the structure of the human AGA precursor and details of the activation mechanism. The WT human (1apy) and bacterial (2gaw) molecules have similar structures (rmsd = 1.12 Å for the backbone of 509 residues) and the residues involved in enzymatic catalysis are conserved. Furthermore, the secondary and tertiary

structures of the bacterial mature AGA (2gaw) (Guo *et al.* 1998) and precursor AGA (9gaf) (Xu *et al.* 1999) are essentially identical (rmsd = 0.58 Å for the backbone of 547 residues). Thus, it is conceivable that the precursor structures of human and bacterial AGA are closely related. Consequently, the activation mechanisms are likely to be similar.

The activation of the AGA precursor and enzymatic catalysis seem to be accomplished by the same amino acids, which makes the interpretation of experimental data complicated. The previous results on human (I) and bacterial (Guan *et al.* 1998) AGA autocatalysis indicate that the mechanism is less efficient than in enzymatic catalysis. Consequently, the analysis of the autocatalytic mechanism using mutagenesis is challenging and the results obtained less accurate. Furthermore, no structural (or other direct) information about the reaction intermediate in Ntn-hydrolase autoproteolysis studies has been obtained. The bacterial AGA precursor structures are very similar except at the linker peptide between the subunits and some neighboring residues of the active site. The nucleophile location and rotamer conformation vary only slightly, but the position and conformation of the preceding residues, including D205, are quite dissimilar. Therefore, to draw detailed conclusions about the residues participating at each activation step would be speculative. Overall, the active site of the bacterial AGA is somewhat wider than the human counterpart. The bacterial enzyme is periplasmic, whereas the human AGA functions in the acidic lysosomes. The human enzyme may have evolved to obtain a different pH optimum which could explain the differences.

10.2.3. Residues with a structural role (III)

The previously published human AGA mutations were evaluated to study whether they could potentially affect the autocatalytic activation and to enhance the structure-function understanding of AGA. In addition, nine novel mutations (T33A, D70A, D200A, D201A, N225A, G226A, K230A, S238A, and G258A) were generated, based on their location in the active site. The effect of the mutations was monitored by immunoprecipitation- and immunofluorescence-based analysis of the expressed polypeptides, and the enzymatic activity and kinetic parameters of AGA were measured. Table 2 summarizes the data.

In the following, only the most interesting structural residues are discussed. Structural comparison of four Ntn-hydrolases has revealed nine fully conserved secondary structure elements (Oinonen and Rouvinen 2000), all located in the β-subunit of the active AGA molecule (Figure 6). In addition to the AGAs of various species (human AGA, 1apy), G258 is also conserved in PRO and GAT, while PGA has alanine at the position. At the position corresponding to G226, only hydrophobic residues are found in other Ntn-hydrolases, while G226 is fully conserved in the AGAs of several species (II).

Residues G258 and T257 form the human AGA oxyanion hole that is utilized in the stabilization of the negatively charged carbonyl oxygen on the transition state during enzyme catalysis (Figure 5). G258A prevented activation and the precursor polypeptide existed in cells in dimeric and monomeric forms. The results indicate that G258 plays a crucial structural role in the activation of the human AGA precursor. The mutant polypeptides are probably unable to fold correctly and are mostly retained in the ER.

G226 resides in a loop at the active site and the substituting aspartate disturbs local folding by eliminating a tight turn of the loop and presenting a larger side chain in the aqueous environment of the substrate-binding funnel. The G226A polypeptides showed a severe folding defect. The main chain nitrogen of G226 (G172) is hydrogen-bonded with the side chain of T224 (T170) in the human and bacterial AGA structures. The side chain of D226 or A226 would create clashes. Possibly, the aspartate side chain could stabilize the structure by creating electrostatic interactions with K230. A lack of such stabilization would create a major folding defect like that observed in the G226A mutant.

Previous data exist for additional mutations, some resulting in human AGU. The hydrogen bond between S72 and T206 is required for efficient enzymatic catalysis (Tikkanen *et al.* 1996a). S72P mutant precursor polypeptides, found in AGU patients (Peltola *et al.* 1996), are not activated intracellularly owing to disturbed packing caused by a change in a loop structure at the active site.

T257A (Tikkanen *et al.* 1996a) and the AGU-causing T257I mutant prevent activation of the precursor. However, the mutagenesis results of T257 are different between human and bacterial AGA (Tikkanen *et al.* 1996a, Liu *et al.* 1998), possibly reflecting different catalytic roles in the enzymes. In some bacterial precursor AGA structures (9gac and 9gaa), T257 (T203) is hydrogen-bonded to the side chain of D205 (D151), while in others (9gaf, 1p4k, and 1p4v) the side chain orientation of D205 is different. The C^{α}-position (~1 Å) and rotamer conformation of T257 also vary. Thus it can not be concluded with certainty that the role of human AGA T257 is to properly orient the side chain D205 with a hydrogen bond. Figure 8 summarizes the roles of all the residues studied.

Construct	Processing	Targeting	Activity % WT	Kinetic data	Problems caused by substitution
WT	(Prec), pro-α, α, β	Lysosomes	100	Normal	
COS-1			≤ 7	K_m ↑↑, V_{max} ↓↓	
WT residues participate in autocatalysis					
D205A/G/S	Prec, α, pro-β	nd	≤ 7	nd	Conformational strain at activation site missing, possible role in autocatalysis
T206A/C	Prec, α, pro-β	nd	≤ 7	nd	Nucleophile defective in catalysis and autocatalysis
[a]T206S	Prec + WT	nd	≤ 7	nd	Nucleophile defective in catalysis, autocatalytic nucleophile weakened
[b]T224A	Prec + WT	nd	≤ 7	nd	T206 destabilized in catalysis, local structural change, possible stabilizing role in autocatalysis
[b]T224S	as WT	nd	21	V_{max} ↓	T206 stabilization affected in catalysis
WT residues with a structural role					
T33A	as WT	Lysosomes	48	K_m ↑	Catalysis affected via local structural change
[a]T33S	as WT	nd	100	nd	
[a]W34F/S	Prec + WT	nd	14	nd	Local structure affected; activation affected possibly via H204
D70A	Prec + WT	Lysosomes	44	K_m ↓, V_{max} ↓	Initial delay in activation, folding affected
[b]S72A	as WT	nd	52	V_{max} ↓	T206 stabilization lacking in catalysis
[c]S72P	Prec, α, pro-β	Lysosomes	30	nd	Disturbed packing, loop structure changed
D200A	Prec + WT	Lysosomes	87	Normal	Folding mildly affected
D201A	Prec + WT	Lysosomes	93	Normal	Folding mildly affected
H204G/[c]S	Prec, pro-α, α, pro-β	nd	30	nd	Conformational strain at activation site decreased, local structural change
N225A	Prec + WT	Lysosomes	45	K_m ↑	Delayed activation, local structural change, may affect S72 and T224
G226A			≤ 7	as COS-1	Hydrophobic residue in Ntn-hydrolases, tight backbone loop eliminated, serious folding defect
G226D	Prec, α, pro-β	Lysosomes	≤ 7	nd	Side chain may stabilize molecule, affects active site structure
K230A	Prec + WT	Lysosomes	86	K_m ↑	Initial delay in activation, affects folding
[b]R234A/K/[c]Q	Prec, α, pro-β	nd	≤ 7	nd	Substrate binding defective, defective folding, destabilizes local structure (and autocatalytic residues)
[b]D237A/S	Prec + WT	nd	≤ 7	nd	Substrate binding defective, delayed folding
S238A	Prec + WT	Lysosomes	40	K_m ↑, V_{max} ↓	Delayed processing, may affect substrate binding via D237 and local structural change
T257I/[b]A	Prec, α, pro-β	Lysosomes	≤ 7	nd	Defective oxyanion hole in catalysis, structural change in active site via autocatalytic residues
[b]T257S	as WT	nd	70	V_{max} ↓	Catalytic efficiency decreased
G258A	Prec, α, pro-β	Lysosomes	≤ 7	as COS-1	Gly partially conserved in Ntn-hydrolases, defective oxyanion hole in catalysis, active site structure defective, delayed folding, mostly retained in the ER, dimerization affected

Table 2. Summary of immunoprecipitation, immunofluorescence, activity measurement, and kinetic data in active site mutations. nd, not determined; [a](Riikonen *et al.* 1995), [b](Tikkanen *et al.* 1996a), [c](Peltola *et al.* 1996).

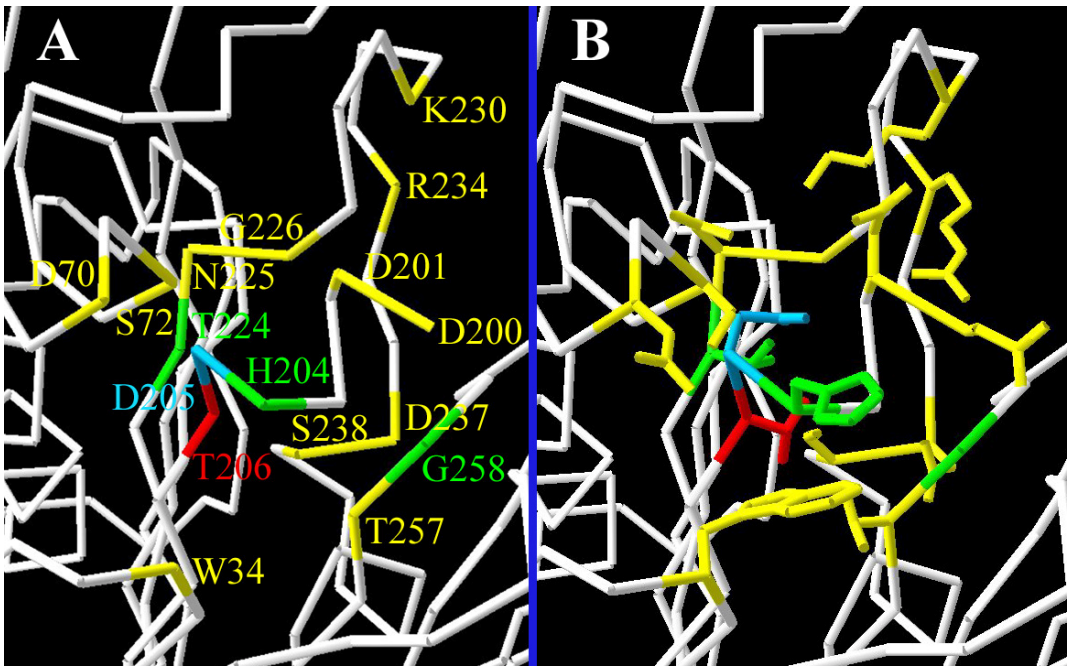

Figure 8. The autocatalytic residues in the WT AGA structure. A) The polypeptide backbone. To show all the residues substituted, D205, H204, G203, R202, D201, and D200 have been built, without energy minimization into an α-helix in this figure. Conformation of the backbone and the side chains of these residues do not correspond to any AGA precursor structure. Native AGA residues are unaffected by the added residues. B) The WT side chains of the residues that affect autocatalysis are shown. T206 is the key residue in enzyme catalysis and the autocatalytic activation. D205 and possibly H204 provide conformational strain at the activation site which may trigger autocatalytic activation. T224 and G258 are important structural residues. The other residues, shown in yellow, affect autocatalysis by providing a structural scaffold for the participating residues.

10.2.4. Peptide bond distortion (I)

Inteins are typical protein structures capable of self-splicing (Perler *et al.* 1994). The structure of the N-terminal cleavage site in the GyrA intein (Klabunde *et al.* 1998) revealed that the polypeptide backbone at the cleavage site was strained because of an energetically unfavorable *cis*-peptide bond. Although Ntn-hydrolases are not related to inteins, the autocatalytic peptide bond cleavage mechanisms have similarities. In AGA, substitutions of the fully conserved residue T206 hindered the processing, suggesting that the side chain of T206 would directly participate in catalysis. Substitutions of the less conserved D205 prevented the processing as well. This could indicate that the proper geometry of D205 would be essential for the autoproteolysis. In line with this observation, H204 could affect the geometry and activation but less efficiently than D205 does. Thus,

prior to the first publication among Ntn-hydrolases to show conformational strain at the autoproteolytic site (Ditzel *et al.* 1998), AGA precursor activation was proposed to be dependent on strain of the polypeptide backbone.

In contrast to the GyrA structure, some bacterial AGA precursor structures contain a high energy distorted *trans* peptide bond between D205 (D151) and the nucleophile T206 (T152) (Xu *et al.* 1999). These residues involved in the N->O acyl rearrangement form an unusual tight turn. The main chain distortions could raise the energy level of the activation site and generate the activation potential for autoproteolytic cleavage (Xu *et al.* 1999). The dimerization of AGA precursors may create polypeptide backbone strain near the scissile peptide bond. As a result, autoproteolysis may be triggered to remove the strain and generate the essential N-terminal threonine (Wang and Guo 2003). Strained splicing junctions are also involved in protein splicing of PI-SceI, which is processed at two distorted *trans* peptide bonds (Poland *et al.* 2000). Much like the GyrA structure, there does not appear to be a good general base in close proximity to C1 at the N-terminal splice junction.

Among the Ntn-hydrolases PRO, GAT, PGA, CA, and GCA, the immediate N-terminal residue of autoproteolytic residue is glycine. The AGA sequence at the processing site is His-Asp-Thr except in plants where it is Ile-Gly-Thr. In addition, inteins (Perler *et al.* 1994), aspartate decarboxylase (Albert *et al.* 1998), and hedgehog protein (Hall *et al.* 1997) have a conserved glycine at the scissile peptide bond. PRO adopts a well-defined γ-turn conformation at the autoproteolytic site (Ditzel *et al.* 1998). The spacer peptide also contains conformational strain on the precursor polypeptide of GCA (Kim *et al.* 2003) and CA (Kim *et al.* 2002) and high-energy state of the spacer peptide has been proposed to trigger the autoproteolytic cleavage in GCA. In addition, the PGA precursor structure indicates non-ideal stereochemistry of the scissile peptide bond (Hewitt *et al.* 2000). Thus although several similarities (overall topology, similar active site residues, conformational strain) exist in Ntn-hydrolase activation, the details of the mechanisms probably diverge, possibly even within the (glyco)asparaginases.

10.2.5. Activation mechanism of human AGA (I, III)

Different bacterial precursor mutants display surprisingly different structures (Xu *et al.* 1999, Qian *et al.* 2003). Backbone and side chain conformation of the residues in the linker peptide between the α- and β-subunits, including D205 (D151), and some active site residues differ significantly between the structures. Therefore, conclusion of the exact role of individual residues at each step of human AGA activation is not possible. As a consequence, the bacterial AGA activation mechanism (Xu *et al.* 1999), which was concluded from the data obtained from three different structures (9gaa, 9gac, 9gaf), may

not represent the best model for human AGA activation. This fact and the human AGA experimental data suggest that the activation mechanism proposed for human AGA (I) remains a valid reaction mechanism for human AGA autoproteolysis.

According to the activation mechanism (Figure 9), the side chain of T206 would play an active role in the processing of AGA. The nucleophilicity of T206 is enhanced by a general base, which is either the side chain of D205 or a water molecule, as discussed later. Thus the general base has a key role in initiating the N->O acyl shift. During the acylation step, the side chain hydroxyl group of T206 forms a hydrogen bond with the carbonyl carbon of D205, and a tetrahedral intermediate is formed. The intermediate is probably stabilized an oxyanion hole. During the deacylation step, a neighboring water molecule probably hydrolyses the ester bond resulting from the N->O shift and resolves the α- and β-subunits of AGA. At this instant, the α-amino and side chain hydroxyl groups of T206 are ready for enzymatic catalysis. This reaction mechanism would require the existence of one base and one acid that assist in the reaction. Unfortunately, without human precursor structure the general base has not been identified in human AGA. Most of the active site residues analyzed do not appear to have a distinct role in the autoproteolytic activation reaction although they are important for the reaction. They surround the active site and form the necessary hydrogen bond network to position and interact correctly with the residues that participate in the reaction.

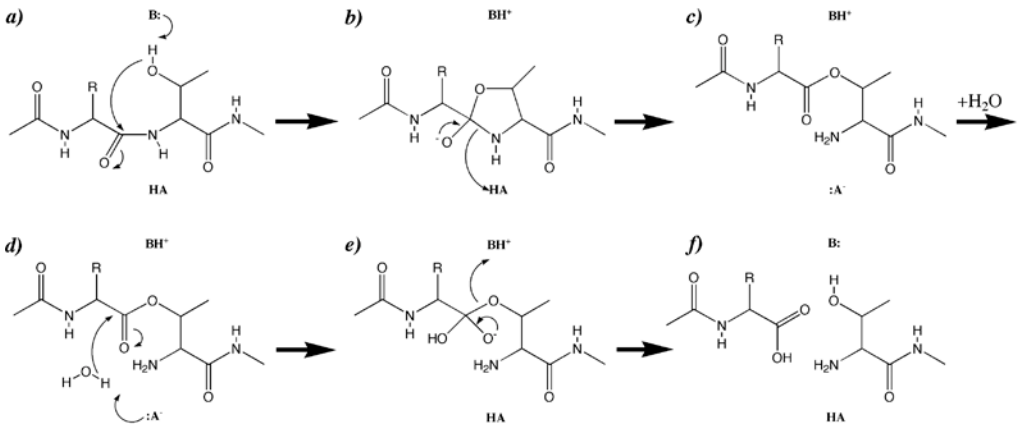

Figure 9. The autocatalytic mechanism of human AGA. The activation is initiated by a general base (A), which receives the hydrogen atom from the side chain of T206 thus enabling a nucleophilic attack on the carbonyl carbon of D205 and resulting in a tetrahedral intermediate (B). C) An N->O acyl shift occurs and a general acid donates a proton to newly formed amino group terminal of T206. D) A water molecule hydrolyses the resulting ester bond and a second tetrahedral intermediate is formed (E). F) The intermediate is resolved and T206 with free amino and side chain groups is ready for enzymatic catalysis.

10.2.6. Autocatalytic activation in Ntn-hydrolases

With T1 as the nucleophile in proteasomal enzyme catalysis, a proton acceptor in the active site is required for activation of the hydroxyl group. A catalytic water molecule, seen in the high-resolution structure of yeast PRO (Groll *et al.* 1997), probably acts as the general base to help activate the T1 hydroxyl group. Several reports of Ntn-hydrolase precursor structures indicate a water molecule as the general base (Ditzel *et al.* 1998, Kim *et al.* 2002, Kim *et al.* 2003, Yoon *et al.* 2003). An acid and a base are also required for the autoproteolytic activation of the AGA precursor (Figure 9). In the case of bacterial precursor AGA, D205 (D151) was proposed to act as the general base that was hydrogen-bonded to the nucleophile in one precursor structure (9gaf). As discussed earlier, the position and conformation of D205 varies with different bacterial AGA precursor mutants. Thus, it is possible that also for human AGA the general base may be a water molecule. In the WT human AGA structure (Oinonen *et al.* 1995), the best candidate water molecule is hydrogen-bonded to T206 and D237. In this case, D205 would only have a role in the formation of main chain distortion, which may trigger the cleavage of the scissile peptide bond. The conformational strain is reduced by the substitution of structurally important residues. These mutants thus have defective autocatalytic machineries that delay cleavage of the precursor. Compared to usual enzymatic catalysis, the autocatalytic cleavage of the peptide bond seems to be less efficient. It is worth emphasizing that in eukaryotic cells the activation of AGA occurs in the ER whereas the enzymatic catalysis takes place in the lysosomes, which could contribute to the lower efficiency of the autocatalytic cleavage when compared to the enzyme catalysis.

CA and GCA processing is very similar to AGA including the removal of the signal sequence and the C-terminal peptide of the α-subunit (Kim *et al.* 2002, Kim *et al.* 2003). However, in AGA the second peptide bond cleavage is performed by a lysosomal protease, whereas in the case of CA and GCA the cleavage is probably performed autocatalytically by the enzyme molecule itself. As in other Ntn-hydrolases, the nucleophilic N-terminal residue of the β-subunit plays an essential role in catalysis and autocatalysis in GCA and CA. Furthermore, the precursor molecule structures are very similar to the native structure of the enzymes. During the activation of the GCA precursor a water molecule is utilized as a general base to deprotonate the hydroxyl group of S170. In addition, a bound water molecule seems to initiate the autoproteolytic activation of CA (Yoon *et al.* 2003).

Ntn-hydrolases have a common structural fold with topological secondary structures. Although human AGA shares sequence similarity with other Ntn-hydrolases only in certain secondary elements, the autoproteolytic activation process shares several features with them (Ditzel *et al.* 1998, Xu *et al.* 1999, Kim *et al.* 2003): 1) The overall structure of the precursor and the active molecule is very similar. 2) The hydroxyl or sulfhydryl group of the nucleophile, which functions both as the catalytic and autocatalytic nucleophile, is oriented

towards the scissile peptide bond. 3) Conformational strain exists in the precursor at the cleavage site. 4) Water molecules are utilized in the activation process. Table 3 summarizes the data on Ntn-hydrolases.

	Human AGA	**Bacterial AGA**	**PRO**	**PGA**	**GAT**	**GCA**	**CA**
Structure	$(\alpha\beta)_2$ hetero-tetramer	$(\alpha\beta)_2$ hetero-tetramer	7 member ring	αβ dimer	homo-tetramer	$(\alpha\beta)_2$ hetero-tetramer	αβ hetero-dimer
Source	human	bacterial	archaea yeast	bacterial	bacterial	bacterial	bacterial
Catalytic function	glycosyl-asparaginase	glycosyl-asparaginase	protease	penicillin acylase	amidotra-nsferase	deacylation of cephalo-sporin	deacylation of cephalo-sporin
Activation	cleavage into subunits	cleavage into subunits	propeptide removal	cleavage into subunits	propeptide removal	cleavage into subunits	cleavage into subunits
WT structure PDB code	1apy, 1apz	1ayy, 2gaw	1pma	1pnk, 1gk9	1gph, 1ecf	1or0	1fm2, 1gk0, 1gk1
Nucleophile	T206	T152	T1	S264	C1	S170	S1β
Activation site	H204-D205-T206	H150-D151-T152	L-2 - G-1 - T1	T262-G263-S264	E-2 - E-1 - C1	Q168-G169-S170	Q168α-G169α-S1β
Oxyanion hole in catalysis	T257, G258	T203?, G204?	L33/G47/ S129	A332, N504	N101, G102	H192?, water?	N244β, V70β
Precursor structure PDB code	nd	9gaa, 9gac, 9gaw, 1p4k, 1p4v	1ryp	1e3a	nd	1oqz	1keh
Proposed oxyanion hole in auto-catalysis	T224?	T170, water	nd	A332?, N504?	N101?, G102?	N413, H192	water, H23β
Proposed general base in auto-catalysis	water/ D205	D151	water	water?	water?/ H70?	water	water
Confor-mational strain	probable	yes	γ-turn	probable	probable	yes	probable
Extra cleavage	Q190-D191, E197-T198	-	-	Y260-P261	-	G160-D161	-

Table 3. Summary of Ntn-hydrolases.

11. Characterization of AGU mutations (II)

The 3D-structure of human AGA made it possible to analyze the structural consequences of AGU mutations. The AGU mutations affect the folding, dimerization, activation, or transport of the AGA molecule and reveal residues crucial for any of these necessary steps. To identify the molecular defect in aspartylglucosaminuria in uncharacterized patients, the patients' genomic DNA was amplified by PCR with *AGA* specific primers. The nucleotide sequences of each exon, exon-intron boundaries, and untranslated regions were determined, and seven previously unknown mutations were identified. The novel mutations are schematically presented in Figure 10 in addition to previously published mutations. AGU mutations span the entire *AGA* cDNA except exon 5. The mutation spectrum is wide with missense and nonsense mutations, deletions, insertions, and splicing defects.

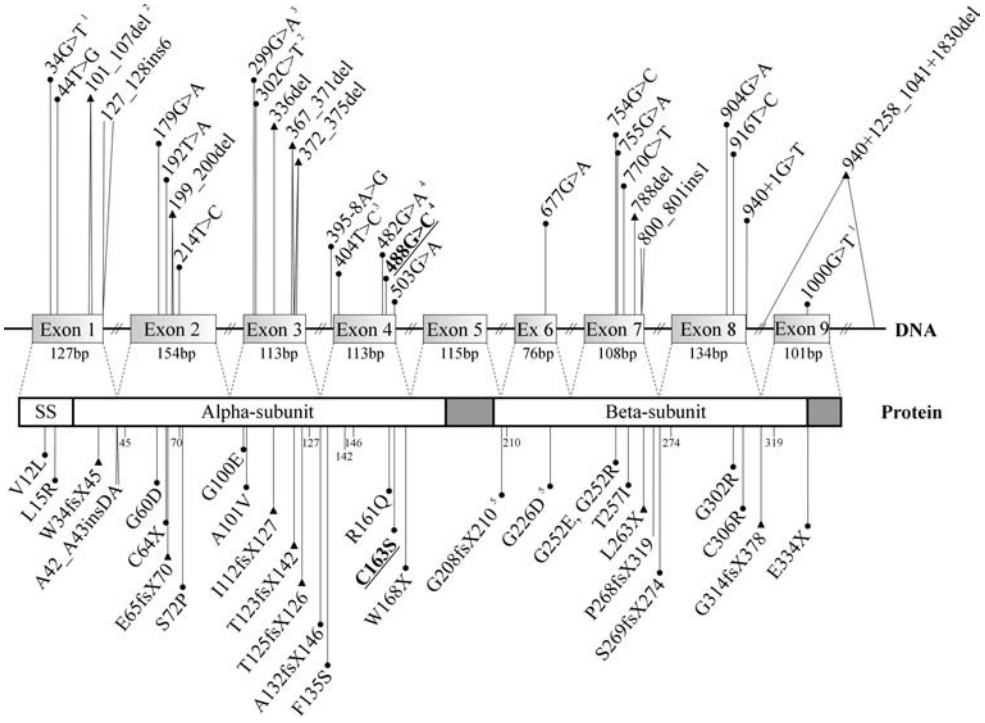

Figure 10. Map of the known AGU mutations. Nucleotide changes are presented above the exons. Corresponding changes at the polypeptide level are presented below the AGA precursor polypeptide. Substitutions are marked with a dot. Deletions are marked with a triangle; the lines connecting the label to the exon illustrate the deleted region. Insertions are marked with two lines, which part at the label end. The disease causing mutation of the AGU_{Fin} major allele is highlighted. [1,4]Concurrent mutations. [2,3]Compound heterozygote mutations. [5]Two different mRNA species in a patient. Copyright (2001) Oxford University Press, UK.

11.1. Consequences of AGU mutations

AGA cDNA was mutagenized to create the new mutation constructs which were expressed. The effects of the mutations on the maturation of the AGA polypeptide were studied using the SVpoly expression vector and transient expression in COS-1 cells. The processing of AGA polypeptides after chase times of 1, 3, and 6 hours was monitored. Among 28 AGU mutations studied, all but V12L revealed problems in intracellular processing and 24 prevented the activation cleavage in the ER. The intracellular distribution of AGA polypeptides was studied in Syrian golden hamster kidney cells (BHK-21) where protein synthesis was stopped for three hours with cycloheximide to allow the proteins to be transported from the ER. Among 28 AGU mutations studied, 22 demonstrated significant problems in the lysosomal targeting of AGA polypeptides.

The mutations are here grouped according to the severity of the effect for the AGA polypeptide, judged based on intracellular maturation, lysosomal targeting, or activity. They are considered to cause mild, moderate, or severe effects. The clinical phenotype of AGU patients is reported to be relatively homogenous (Aula *et al.* 2001). However, the clinical details of the progression or severity of symptoms are unavailable and at this point it is not possible to reliably analyze if the severity of mutations at the polypeptide level is reflected in the clinical outcome of patients.

11.1.1. Mutations with mild effects

AGU mutations that produce intracellularly stabile precursor polypeptides and at least a fraction of normal subunits are here defined as mild. Intracellularly, these AGA polypeptides are seen in the Golgi compartment and lysosomes in addition to the ER. The polypeptides may also have detectable enzyme activity.

The unpublished L15R mutation of the signal sequence is mild with 30% activity compared to the WT. Signal sequences typically contain many hydrophobic residues. The mutation is located in a long stretch of hydrophobic residues and thus it could affect polypeptide synthesis by delaying transport of AGA through the translocon complex. The other mutation of the signal peptide, V12L, had no effect on the AGA polypeptide.

S72P disturbs packing of the molecule by introducing proline into a loop in the active site resulting in abnormal maturation of polypeptides. Also stabilization of T206 during catalysis is prevented.

F135S mutation affects dimerization and subsequently activation by destabilizing the loop which contains H124 (see 10.1).

The tight backbone turn at G226 would be absent in the G226D substitution, since according to the Ramachandran plot such backbone torsion angles can only be allowed for glycine. The substitution might cause spatial problems and in any case the long side chain of aspartic acid would occupy the active site preventing normal processing.

There are only a few AGU-causing active site mutations of AGA. So far, the only mutations of residues directly involved in catalysis are the homozygous S72P and the novel heterozygous mutation T257I. T257 and G258 form the oxyanion hole that stabilizes a covalent enzyme-substrate transition state (Tikkanen *et al.* 1996a). The larger isoleucine is unable to replace T257 during enzymatic catalysis.

11.1.2. Mutations with moderate effects

AGU mutations are here classified as moderate when they produce inactive precursor polypeptides, which remain unprocessed. These precursor polypeptides are retained in the ER and the majority of polypeptides face rapid degradation.

The A42_A32insDA mutation disrupts an α-helix. The insertion most probably disturbs local folding in the ER, since the structural environment of A43 is restricted and the mutation changes the residue order of the helix.

Similar outcome resulted from G60D mutation since the side chain of aspartic acid is excessively large for local structure and results in misfolding of the polypeptide. G100E and A101V disturb the three-dimensional structure of AGA because of space limitations and probably also affect dimerization of precursors.

C163S substitution is responsible for defective folding and premature degradation in the AGU_{Fin} major mutation, R161Q+C163S. C163S causes structural problems in addition to preventing the formation of the disulfide bridge between C163 and C179 resulting in destabilization of the loop structure that provides contacts between the two αβ dimers of AGA. The reason the mutated polypeptides remain in the ER as monomeric precursor polypeptides is evidently that the initial folding is defective and dimerization is prevented (Riikonen *et al.* 1996). On the other hand, C163 is located very near to K177 and Y178, which form one of the three phosphotransferase recognition sites and the substitution also prevents phosphorylation by altering local folding (Tikkanen *et al.* 1997).

The unstable precursor molecules produced by the novel mutations G252E and G252R can be explained by the fact that according to the Ramachandran plot the backbone torsion angles at 252 are only allowed for glycine. G302R, a neighbor of the oxyanion hole residue T257, also made the precursor unstable.

The mutations with an abnormal C-terminus of the β-subunit produced unstable precursor polypeptides. Quite surprisingly, some precursor polypeptide could be observed with PAGE and immunofluorescence analyses (see 10.1).

11.1.3. Mutations with severe effects

AGU mutations, which make the AGA polypeptide remarkably unstable, are here defined as severe. Trace amounts of mutant polypeptides in this category are seen after a 6 h chase, but they do not represent normal AGA precursor or subunits. AGA polypeptide is undetectable in immunofluorescence analysis of transfected cells even with AGA antibodies against the denatured subunits. The mutant polypeptides are severely shortened by premature stop codons of the *AGA* gene.

The only severe mutation with a point mutation is C306R. Replacement of the disulfide bond forming C306 with arginine results in total misfolding and rapid degradation of the synthesized polypeptides. The substitution disturbs the packing of the molecule. This is the determining factor in the misfolding, since C286S and C306S polypeptides were normally processed and active demonstrating that the disulfide bridge between C306 and C286 is not crucial for the early folding and stability of AGA (Riikonen *et al.* 1996).

Table 4 lists the patient mutations and Figure 11 shows the locations of AGU point mutations in the 3D-structure of the WT molecule.

Construct	Processing	Targeting	Activity % WT	Comments
WT	(Prec), pro-α, α, β	Lysosomes	100	
COS-1			≤ 7	
Mild mutations				
[a]**V12L**	(Prec), pro-α, α, β	Lysosomes	100	Like WT
L15R	Prec, pro-α, α, β	ER, Lysosomes	30	May affect synthesis
[b]S72P	Prec, α, pro-β	ER, Lysosomes	30	Restricted space, affects packing, defective active site
F135S	Prec, pro-α, α, β	ER, Lysosomes	12	Destabilization,affects dimerization
G226D	Prec, α, pro-β	ER, Lysosomes	≤ 7	Tight backbone turn missing
T257I	Prec, α, pro-β	ER, Lysosomes	≤ 7	Defective active site, restricted space
Moderate mutations				
A42_A43insDA	Prec	ER-trapped	≤ 7	Disrupts α-helix, affects packing
G60D	Prec	ER-trapped	≤ 7	Restricted space
G100E	Prec	ER-trapped	≤ 7	Restricted space
A101V	Prec	ER-trapped	≤ 7	Restricted space, affects dimerization
R161Q+C163S	Prec	ER-trapped	≤ 7	C163S: Restricted space, prevents SS-bridge formation and dimerization, defective targeting
G252E	Prec	ER-trapped	≤ 7	Backbone torsion angle disallowed
G252R	[e]Prec	ER-trapped	≤ 7	Backbone torsion angle disallowed
P268fsX319	[e]Prec	ER-trapped	≤ 7	C-terminus of β-subunit altered
[c]S269fsX274	[e]Prec	ER-trapped	≤ 7	C-terminus of β-subunit altered
G302R	Prec	ER-trapped	≤ 7	Affects active site
G314fsX378	[f]Prec	ER-trapped	≤ 7	C-terminus of β-subunit altered
[a]E334X	[e]Prec	ER-trapped	≤ 7	C-terminus of β-subunit missing
Severe mutations				
[d]W34fsX45			≤ 7	Short abnormal polypeptide
[b]C64X			≤ 7	Short abnormal polypeptide, partial read-through into precursor
E65fsX70			≤ 7	Short abnormal polypeptide
I112fsX127			≤ 7	Short abnormal polypeptide
[d]T123fsX142			≤ 7	Short abnormal polypeptide
T125fsX126			≤ 7	Short abnormal polypeptide
A132fsX146			≤ 7	Short abnormal polypeptide
W168X			≤ 7	Short abnormal polypeptide
L263X			≤ 7	C-terminus of β-subunit missing
C306R			≤ 7	Restricted space, affects packing

Table 4. Summary of immunoprecipitation, immunofluorescence, and activity measurement data in patient mutations. The novel mutations are shown in boldface type. [a]Concurrent mutations; [b](Peltola *et al*. 1996), [c](Fisher and Aronson 1991), [d](Park *et al*. 1993); [e]Truncated precursor; [f]Elongated precursor.

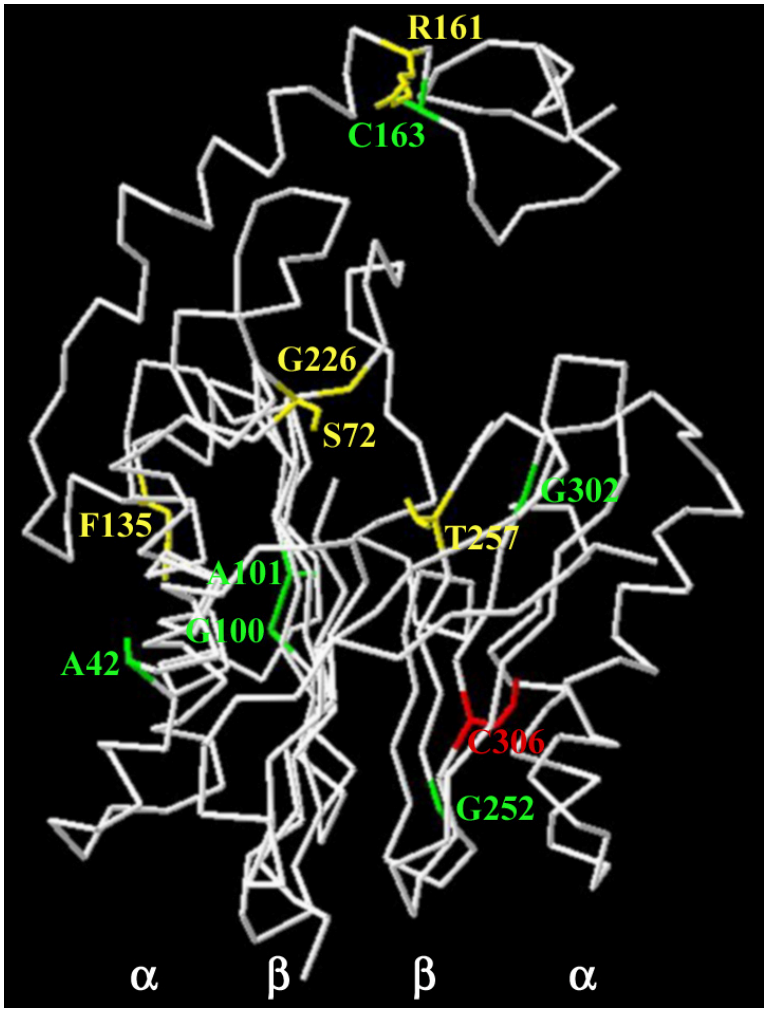

Figure 11. The locations of AGU point mutations. Only one α- and β-subunit of the AGA molecule is shown. The mild mutations are shown in yellow, moderate mutations are green, and the severe mutations are red. The α-helical and β-sheet layers typical for N-hydrolases are indicated at the bottom.

12. Expression and purification of recombinant human AGA precursor

12.1. Expression of recombinant human AGA precursor

Many post-translational modifications performed by higher eukaryotic cells are applied to proteins produced in yeast expression systems. These include processing of the signal sequences, disulfide bridge formation, and O- and N-linked glycosylation. Thus folding of recombinant proteins is generally very native-like. However, glycosylation is not identical to mammalian systems. In lower eukaryotes, O-linked oligosaccharides contain only mannose residues and N-glycosylation is also somewhat different from higher eukaryotes (Cregg *et al.* 2000). The two major yeast systems are *Saccharomyces cerevisiae* (Eckart and Bussineau 1996) and *Pichia pastoris* (Cregg *et al.* 2000). They lack detectable endotoxins and are generally regarded as safe. High expression levels of therapeutic proteins (up to 1.5 g/l) have been achieved using fermentation of *P. pastoris* (Wang *et al.* 2001), which is capable of metabolizing methanol as its sole carbon source.

The recombinant human AGA was produced in the yeast *P. pastoris*. The mutated *AGA* gene encoding the precursor form of the molecule was cloned in several tandem repeats into the yeast genome. The most efficiently producing clone was selected from several different vector and yeast strain combinations. The best clone utilized methanol to promote AGA expression which was approximately at the level of 14 mg/l. The precursor molecule was secreted into the culture medium using the signal sequence of AGA. Several batches of precursor expression were performed in baffled flasks. The production of the precursor using fermentation was also attempted. However, all precursor molecules were cut into two polypeptides during fermentation, which did not occur during the flask expression. During the expression, EDTA was used to inhibit metalloproteases and glycine was added to stabilize the precursor molecule.

12.2. Purification and crystallization of human AGA precursor

Fusion protein vectors are increasingly popular. A fusion protein with a tag of known size and biological function can simplify the isolation, purification, and detection of the target protein and enhance the expression level and secretion. The tags can be helpful for stabilizing fused proteins and enhance their solubility. Expression tags were not used for the AGA precursor because the use of the short hydrophilic FLAG-epitope tag (Chubet and Brizzard 1996) or the HA-tag (Murray *et al.* 1995) in the termini or the loops of the WT AGA molecule was not successful (Kyttälä, A., personal communication). The tagged AGA molecules were either unstable or the tag was cleaved off by proteases. Thus it was necessary to utilize several liquid chromatography methods in the precursor purification scheme (Figure 12). The scheme utilized for the purification of the active enzyme was unsuccessful for the precursor and therefore a new scheme was developed.

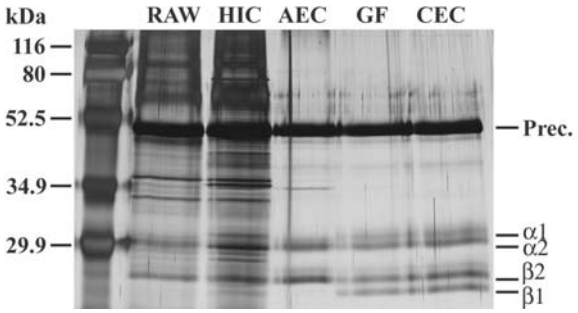

Figure 12. The AGA precursor after liquid chromatography purification steps. In the silver stained PAGE, shown are the raw material after AGA expression (RAW), and the sample after hydrophobic interaction chromatography (HIC), anion exchange chromatography (AEC), gel filtration (GF), and cation exchange chromatography (CEC). The precursor is cut into abnormal subunits during purification. The α1- and β1-subunits result from a cut between E197 and T198 while the α2- and β2-subunits result from a cut between Q190 and D191.

During purification, the AGA polypeptides were observed to aggregate when exposed to light. The aggregation was not dependent on temperature. Thus, the exposure of AGA samples to light during purification was minimized. The precursor molecule was slowly cut into subunits in spite of the use of protease inhibitor cocktail during the purification. N-terminal sequencing revealed that the signal sequence had been normally cut out from the α-subunit but the β-subunits contained two abnormal N-termini (Kalkkinen, N., personal communication). The precursor had been cut 8 and 15 residues before the normal activation site between D205 and T206 (Figure 6). The cut at Q190-D191 occurred in the early steps of the purification, while the cut between E197 and T198 is visible only after anion exchange chromatography. Based on the bacterial precursor structure, the linker

peptide between the α- and β-subunits at these cut sites is located on the surface of the molecule and thus it is probable that yeast proteases are responsible for the hydrolysis of the peptide bonds. Transient expression in COS-1 cells with these cut site residues mutagenized individually and combined did not show aberrant precursor processing but these constructs were not further expressed in *P. pastoris*. The rate of cleavage was dependent on temperature and thus the purification steps were performed with cooled buffers. Table 5 summarizes the purification parameters for the AGA precursor after each purification step during the single largest precursor purification batch. The amount of purified precursor AGA totalled 686 μg (1.37% of the total protein). Overall, several milligrams of the AGA precursor was purified.

Step	Mass total protein, before (μg)	Mass total protein, after (μg)	Recovery total protein (%)	Relative AGA amount (%)	Relative AGA precursor amount (%)
Initial concentration	nd	nd	nd	22	15
HIC	50214	32137	64.0	24	17
AEC	25334	3536	14.0	74	52
GF	3536	2493	70.5	84	54
CEC	2100	937	44.6	84	46

Table 5. Purification of recombinant human AGA precursor from *P. pastoris*. The values have been approximated by scanning and quantitating bands on SDS-PAGE using the Scion Image –program (Scion Corporation).

The separation of the molecules with different β-subunits proved very difficult. Two-dimensional SDS PAGE was used to evaluate the purity and properties of the purified AGA sample. At the molecular weight corresponding to the precursor, three different dots were observed. The two major dots had a 0.16 pH unit difference. In addition, the α- and β-subunits were observed. The different precursor molecules probably represented different glycosylation forms since the mass-spectrometric spectrum of the precursor showed several peaks whose mass-differences correspond to the mass of mannose (Kalkkinen, N., personal communication). The range of the sugar residues was 18-24 and the tallest peaks were at 19-20 residues. Thus in one N-glycosylation site the average oligosaccharide chain contained 2 GlcNAc residues and approximately 8 mannose residues, which corresponds well with the mammalian and the *P. pastoris* high-mannose glycosylation pattern (Miele *et al.* 1997a, Miele *et al.* 1997b). Upon cleavage of the oligosaccharide chains with Endoglycosidase H and Peptide:N-Glycosidase F, the resulting polypeptides were of equal size, further confirming that the oligosaccharide chains of the recombinant AGA precursor were high-mannose-type. The distribution of the oligosaccharide chains or the possible phosphorylation of the recombinant AGA precursor was not analyzed.

The purified AGA precursor samples were delivered to the University of Joensuu, where the crystallization experiments are conducted by Dr. Carita Oinonen. First small hollow cylindrical crystals were obtained and later larger crystals with a different form without a hollow center have been obtained (Figure 13). The best partial data set collected so far from a crystal reached 4.0 Å resolution, which is not sufficient enough for high-detail structure determination. Further crystallization experiments are still ongoing.

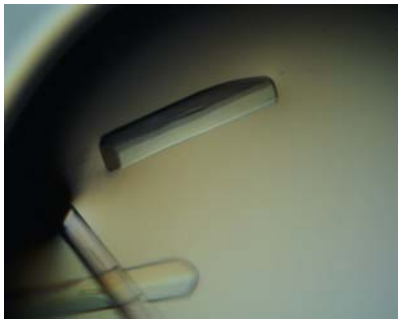

Figure 13. Human AGA precursor crystals.

CONCLUDING REMARKS

AGU was identified in 1968. Since the initial observations, excretion of glycoasparagines into urine and lysosomal accumulation of the substrates of AGA, the AGA enzyme has become one of the best characterized lysosomal enzymes. The AGU mutations have provided clues about structurally and functionally important residues. Prior to this study, the role of disulfide bonds and oligosaccharide chains, and the expression, intracellular processing, lysosomal transport, the three-dimensional structure, and the enzymatic mechanism of the AGA had been characterized in detail. Moreover, initial characterization of the dimerization and activation had been carried out. In addition, using the AGU mouse model therapeutic approaches to treat AGU had been initiated.

This study has elucidated the critical steps in the early stages of AGA maturation in the ER. The activation of newly synthesized AGA precursors requires proper folding and dimerization of the precursor polypeptides and numerous AGU mutations deliver their consequences by disturbing this process. The dimerization was found to be dependent on three regions of amino acids at the dimer interface of the AGA molecule. The activation of human AGA precursors was found to occur autocatalytically. The side chain of T206 was demonstrated to have a central role in the autocatalysis. In addition, the conformation of the polypeptide backbone at the activation site was suggested to be strained. An activation mechanism for AGA autocatalysis was proposed. Later, the structural similarity of the endogenous bacterial AGA precursor and the active enzyme, published by others, also demonstrated that the activation of AGA occurs autocatalytically and conformational strain was found in the activation site. However, the bacterial activation mechanism still contained significant amount of speculation concerning the roles of individual residues.

More active site residues were explored to confirm the activation mechanism of human AGA. The results strongly suggested that the previous mechanism was still valid. Without the human AGA precursor structure the detailed roles of the residues involved in activation still remained unresolved apart from T206 and D205. Several structurally important residues were identified which contribute to the activation by stabilizing and interacting with the core residues of activation.

In addition, seven novel AGU mutations were found and characterized along with previously found mutations. The results provided evidence that structural changes caused by missense mutations are the most common cause of AGU disease. They result in folding problems and destabilization of AGA polypeptides by introducing large residues to limited spaces in three-dimensional structure, by affecting active site residues, by creating electrostatic disturbances, or by preventing disulfide bond formation. The predicted and observed consequences of AGA mutations were in good agreement.

This molecular characterization of activation site mutations and AGU mutations contributes to the expanding body of knowledge pertaining to AGU disease pathogenesis. After molecular characterization of the consequences of these mutations, details of AGA dimerization and activation are now better defined and a more comprehensive structure-function understanding of AGA has been obtained.

Three-dimensional structure of the human AGA precursor would provide a valuable tool that could be used to engineer drugs for AGU treatment. Additionally, enzyme replacement and gene therapy may prove to be efficient in AGU treatment, and future challenges in AGU research include further development of therapeutic methods to alleviate the suffering of AGU patients and their families. This study provides some data which can be utilized to reach the goal. While efficient corrective methods are unavailable, the ethical aspects of genetic counseling and testing should be carefully considered.

ACKNOWLEDGEMENTS

This study was carried out between years 1996-2003 in the Department of Molecular Medicine at the National Public Health Institute in Helsinki. I wish to thank the former head of the Institute, Professor Jussi Huttunen, and the new director Professor Pekka Puska, for providing the excellent research facilities.

I am grateful to my supervisor, Professor Leena Peltonen-Palotie. She provided an excellent research opportunity for me by introducing the fascinating world of proteins. I have been privileged to have an enthusiastic and determined supervisor who inspires young scientists with her tremendous drive and vision.

Docent Anu Jalanko, my unofficial supervisor, always has had time to answer questions and has created a friendly and supportive atmosphere in the lab. I truly appreciate her help and guidance in various matters which range from car rides during public traffic strikes to creating stabile CHO-cell lines.

Professor Ole Kristian Tollersrud and Docent Marc Baumann are gratefully acknowledged for reviewing the manuscript for this thesis and for helpful suggestions. Jean Margaret Perttunen is thanked for revising the language of the thesis.

I have been lucky to work with very talented scientists. Professor Juha Rouvinen (University of Joensuu) is warmly thanked for his input in AGA research. His vast expertise in protein chemistry and enthusiastic attitude carried this project all along. Dr. Carita Oinonen is especially thanked for the crystallization of AGA samples. The task has not been easy but she has managed to crystallize the sometimes not-so-large precursor batches. Juha and Carita are also thanked for their help in preparing manuscripts and for their good sense of humour and Carelian hospitality. My visits to Joensuu have always been a pleasure and an educative lesson on structural biology. The past member of the AGU team, Dr. Ritva “Tiku” Tikkanen is greatly acknowledged for tutoring me during the first summer in LMGO. Dr. Minna Laine is warmly thanked for her help and guidance in the lab and travel company in Barcelona and Åre. Tiku and Minna are also thanked for great co-operation.

Other members of the AGU team, Aija Kyttälä, Kai Tenhunen, Annukka Uusitalo, Annina Lyly, Carina von Schantz, Salli Virta, and Juha Korhonen are also thanked: Kaitsu for entertaining lunch breaks and jokes, Aija, Annukka and Salli for help in various matters, Annina for participating in AGA precursor purification, Tintti for patient DNA sequencing, and Juha for help in native AGA purification. It has been a pleasure working with all of you.

Paula Hakala, Tuula Manninen, Ritva Timonen, and Anne Nyberg are warmly thanked for their excellent technical assistance. Without your skillful participation, I would probably have stress injuries due to pipetting.

Dr. Nisse Kalkkinen is warmly thanked for his advice on protein purification and for allowing me to use the equipment in his laboratory. In addition to Nisse, Dr. Jari Helin, and Dr. Leena Valmu are thanked for performing mass-spectrometric analyses of AGA. Keijo Virta is thanked for help on AGA purification and for nice discussions.

Teemu Perheentupa, Juha Saharinen, Tero Hiekkalinna, Juri Ahokas, Aki Suomalainen, Jari Raikko and Martti Sariola are thanked for keeping the computers and programs up and running during the years. Teemu has also made the work and the parties enjoyable with his humour.

The secretaries of our department, Sari Kivikko, Sari Mustala, Tuija Svahnbäck, and Helena Knuuttila are thanked for their expert help.

Warm thanks go to the patients and families participating in this study and to the collaborators providing me with the samples of the patients: Rita Barone, Catania, Italy; Alan Cooper, Manchester, UK; Paola Rossi and Margherita Santucci, Bologna, Italy; and Pertti Aula, Helsinki, who also was consulted as a clinician.

This study was financially supported by the Academy of Finland (Center of Excellence in Disease Genetics), the Ulla Hjelt Fund of the Pediatric Foundation, the Sigrid Juselius Foundation, the Rinnekoti Foundation, the Finnish Cultural Foundation, and the Maud Kuistila Foundation.

I want to thank the seniors in our department: Ismo “Iski” Ulmanen for his valuable insights and help on various matters and for his humour; Matti Jauhiainen for his willingness to answer my questions, for finding solutions to problems, and for his never-ending good spirit and smile; Outi Kopra for her friendly always-ready-to-help attitude; Irma Järvelä for car rides and discussions; Marjo Kestilä for her interest in the thesis project.

Every past and present people in the department are warmly thanked for friendliness and for good company. Juha Isosomppi is especially thanked for discussions in a wide perspective and for willingness to help on various matters; Tarja Salonen for making the days in Biomedicum bright and happy; Nina Aula for spreading positive energy. Juha, Tarska, and Naula are also thanked for advice on the preparation of this thesis. The new-comers of the writing room, Anna Kiialainen and Heli Honkala, are thanked for keeping up the good humour. Laura Ahtiainen is thanked for friendship and help with adobe-programs. Pekka Ellonen and Heli Keränen are thanked for excellent DNA sequencing service. Jari Metso is thanked for help with the FPLC-machinery and columns. The present members of the lab, Gilberto Duran, Nabil Enattah, Maria Halonen, Henna Haravuori, Ville Holmberg, Tanja Ilmarinen, Ville Karhu, Mikko Kuokkanen, Mira Kyttälä, Markku Lehto, Kaisu Luiro, Katri Miettinen, Niklas Pakkasjärvi, Markus Perola, Seija Puomilahti, Nora Pöntynen, Heli Rasinperä, Taina Rüppel, Joni Turunen, Teppo Varilo, and Tero Ylisaukko-oja are thanked for good company and help in various issues. The past members of the lab, Petra Eskelin, Hannele Kangas, Lasse Lönnqvist, and Kaisu Nikali are thanked for many cheerful moments. In addition, Jyrki Kaukonen is thanked for tough squash matches, and Juha Paloneva and Tuomas Klockars for sharing fly fishing trips and tips (Britton *et al.* 1998). Liina Lonka and Tarja Joensuu, although not from the KTL lab, have nevertheless helped me and are thanked also for the good spirit.

I wish to thank all my dear friends who have been a very important part of my life. Especially Pekka, Mika, and Lasse are thanked for mental and physical coaching for life and for unforgettable moments. Also Esa and Susanna, Jarzu and Karo, Pikku-Jussi and Hantta, Jani K. and Merike, Iso-Jussi and Jaana, Lare, Helmisen Mika, Henri and Merja, Ossi and Anu, Tommi and Päivi, Hanna and Juke, Marja, Salla and Teemu, Sari and Matt, Päivi, Sirke, and Esa are warmly thanked for shared occasions and for being there.

The love and support provided by my family cannot be thanked enough. Isä, Äiti, Mari, and Mira, you have made everything possible. I wish also to thank my brothers-in-law, VP and Mikko, and little Emppu, for bringing joy to life. Sirpa, your love and company means the world to me.

Helsinki, December 2003 Jani

REFERENCES

Albert, A., Dhanaraj, V., Genschel, U., Khan, G., Ramjee, M. K., Pulido, R., Sibanda, B. L., von Delft, F., Witty, M., Blundell, T. L., Smith, A. G. and Abell, C. (1998) Crystal structure of aspartate decarboxylase at 2.2 A resolution provides evidence for an ester in protein self-processing. *Nat Struct Biol*, **5**, 289-93.

Alkema, W. B., Hensgens, C. M., Kroezinga, E. H., de Vries, E., Floris, R., van der Laan, J. M., Dijkstra, B. W. and Janssen, D. B. (2000) Characterization of the beta-lactam binding site of penicillin acylase of *Escherichia coli* by structural and site-directed mutagenesis studies. *Protein Eng*, **13**, 857-63.

Alkema, W. B., Prins, A. K., de Vries, E. and Janssen, D. B. (2002) Role of αArg145 and βArg263 in the active site of penicillin acylase of *Escherichia coli*. *Biochem J*, **365**, 303-9.

Alkhayat, A. H., Kraemer, S. A., Leipprandt, J. R., Macek, M., Kleijer, W. J. and Friderici, K. H. (1998) Human beta-mannosidase cDNA characterization and first identification of a mutation associated with human beta-mannosidosis. *Hum Mol Genet*, **7**, 75-83.

Aridor, M. and Hannan, L. A. (2000) Traffic jam: a compendium of human diseases that affect intracellular transport processes. *Traffic*, **1**, 836-51.

Aridor, M. and Hannan, L. A. (2002) Traffic jams II: an update of diseases of intracellular transport. *Traffic*, **3**, 781-90.

Aronson, N. N., Jr. and Kuranda, M. J. (1989) Lysosomal degradation of Asn-linked glycoproteins. *FASEB J*, **3**, 2615-22.

Arvio, M., Autio, S. and Louhiala, P. (1993a) Early clinical symptoms and incidence of aspartylglucosaminuria in Finland. *Acta Paediatr*, **82**, 587-9.

Arvio, M., Oksanen, V., Autio, S., Gaily, E. and Sainio, K. (1993b) Epileptic seizures in aspartylglucosaminuria: a common disorder. *Acta Neurol Scand*, **87**, 342-4.

Arvio, M., Laiho, K., Kauppi, M., Peippo, M., Leino, P., Kautiainen, H., Kaipiainen-Seppänen, O. and Mononen, I. (2002) Carriers of the aspartylglucosaminuria genetic mutation and chronic arthritis. *Ann Rheum Dis*, **61**, 180-1.

Arvio, M. A., Rapola, J. M. and Pelkonen, P. M. (1998) Chronic arthritis in patients with aspartylglucosaminuria. *J Rheumatol*, **25**, 1131-4.

Arvio, P. and Arvio, M. (2002) Progressive nature of aspartylglucosaminuria. *Acta Paediatr*, **91**, 255-7.

Aula, P., Astrin, K. H., Francke, U. and Desnick, R. J. (1984) Assignment of the structural gene encoding human aspartylglucosaminidase to the long arm of chromosome 4 (4q21 - 4qter). *Am J Hum Genet*, **36**, 1215-24.

Aula, P., Jalanko, A. and Peltonen, L. (2001) Aspartylglucosaminuria. In *The Metabolic & Molecular Bases of Inherited Disease*, **3**, 3535-50, McGraw-Hill, New York

Autio, S. (1972) Aspartylglucosaminuria. Analysis of thirty-four patients. *J Ment Defic Res Monogr Ser I*.

Autio, S., Visakorpi, J. K. and Järvinen, H. (1973) Aspartylglycosaminuria (AGU). Further aspects on its clinical picture, mode of inheritance and epidemiology based on a series of 57 patients. *Ann Clin Res*, **5**, 149-55.

Autti, T., Raininko, R., Haltia, M., Lauronen, L., Vanhanen, S. L., Salonen, O., Aronen, H. J., Wirtavuori, K. and Santavuori, P. (1997) Aspartylglucosaminuria: radiologic course of the disease with histopathologic correlation. *J Child Neurol*, **12**, 369-75.

Bainton, D. F. (1981) The discovery of lysosomes. *J Cell Biol*, **91**, 66-76.

Baldwin, R. L. (2002) A new perspective on unfolded proteins. *Adv Protein Chem*, **62**, 361-7.

Bao, M., Booth, J. L., Elmendorf, B. J. and Canfield, W. M. (1996a) Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. *J Biol Chem*, **271**, 31437-45.

Bao, M., Elmendorf, B. J., Booth, J. L., Drake, R. R. and Canfield, W. M. (1996b) Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. II. Enzymatic characterization and identification of the catalytic subunit. *J Biol Chem*, **271**, 31446-51.

Barlowe, C. (2002) COPII-dependent transport from the endoplasmic reticulum. *Curr Opin Cell Biol*, **14**, 417-22.

Baumann, M., Peltonen, L., Aula, P. and Kalkkinen, N. (1989) Isolation of a human hepatic 60 kDa aspartylglucosaminidase consisting of three non-identical polypeptides. *Biochem J*, **262**, 189-94.

Beckmann, R. P., Mizzen, L. E. and Welch, W. J. (1990) Interaction of Hsp 70 with newly synthesized proteins: implications for protein folding and assembly. *Science*, **248**, 850-4.

Blanco, F. J., Serrano, L. and Forman-Kay, J. D. (1998) High populations of non-native structures in the denatured state are compatible with the formation of the native folded state. *J Mol Biol*, **284**, 1153-64.

Blount, P. and Merlie, J. P. (1991) BIP associates with newly synthesized subunits of the mouse muscle nicotinic receptor. *J Cell Biol*, **113**, 1125-32.

Braakman, I., Helenius, J. and Helenius, A. (1992) Role of ATP and disulphide bonds during protein folding in the endoplasmic reticulum. *Nature*, **356**, 260-2.

Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C., Smith, J. L., Tomchick, D. R. and Murzin, A. G. (1995) A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature*, **378**, 416-9.

Britton, B. J., Evans, J. G. and Potter, J. M. (1998) Does the fly matter? the CRACKPOT study in evidence based trout fishing. The Collaborative Randomised and Controlled Kennet Piscatorial Options Trail (CRACKPOT) Investigators. *BMJ*, **317**, 1678-80.

Bruggink, A., Roos, E. R. and De Vroom, E. (1998) Penicillin acylase in the industrial production of β-lactam antibiotics. *Org Proc Res Dev*, **2**, 128-33.

Buchner, J. (1996) Supervising the fold: functional principles of molecular chaperones. *FASEB J*, **10**, 10-9.

Busca, R., Pujana, M. A., Pognonec, P., Auwerx, J., Deeb, S. S., Reina, M. and Vilaro, S. (1995) Absence of N-glycosylation at asparagine 43 in human lipoprotein lipase induces its accumulation in the rough endoplasmic reticulum and alters this cellular compartment. *J Lipid Res*, **36**, 939-51.

Bustin, M. and Conway-Jacobs, A. (1971) Intramolecular activation of porcine pepsinogen. *J Biol Chem*, **246**, 615-20.

Cantz, M. and Ulrich-Bott, B. (1990) Disorders of glycoprotein degradation. *J Inherit Metab Dis*, **13**, 523-37.

Chen, W., Helenius, J., Braakman, I. and Helenius, A. (1995) Cotranslational folding and calnexin binding during glycoprotein synthesis. *Proc Natl Acad Sci U S A*, **92**, 6229-33.

Choi, K. S., Kim, J. A. and Kang, H. S. (1992) Effects of site-directed mutations on processing and activities of penicillin G acylase from *Escherichia coli* ATCC 11105. *J Bacteriol*, **174**, 6270-6.

Chothia, C. (1975) Structural invariants in protein folding. *Nature*, **254**, 304-8.

Chubet, R. G. and Brizzard, B. L. (1996) Vectors for expression and secretion of FLAG epitope-tagged proteins in mammalian cells. *Biotechniques*, **20**, 136-41.

Conchie, J. and Strachan, I. (1969) Distribution, purification and properties of 1-aspartamido-beta-N-acetylglucosamine amidohydrolase. *Biochem J*, **115**, 709-15.

Coulter-Mackie, M. B. (1999) A novel exonic mutation in the aspartylglucosaminidase gene causes exon skipping. *J Inherit Metab Dis*, **22**, 682-3.

Cregg, J. M., Cereghino, J. L., Shi, J. and Higgins, D. R. (2000) Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol*, **16**, 23-52.

Cui, T., Liao, P. H., Guan, C. and Guo, H. C. (1999) Purification and crystallization of precursors and autoprocessed enzymes of *Flavobacterium* glycosylasparaginase: an N-terminal nucleophile hydrolase. *Acta Crystallogr D Biol Crystallogr*, **55**, 1961-4.

Daggett, V. and Fersht, A. (2003) The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol*, **4**, 497-502.

Dahms, N. M. and Hancock, M. K. (2002) P-type lectins. *Biochim Biophys Acta*, **1572**, 317-40.

de Duve, C., Pressman, B. C., Gianetto, R., Wattiaux, R. and Appelmans, F. (1955) Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat liver tissue. *Biochem J*, **60**, 604-17.

de Duve, C. (1963) The lysosome concept. In *Lysosomes*, 1-35, Ciba Found. Symposium, London

de Duve, C. (1983) Lysosomes revisited. *Eur J Biochem*, **137**, 391-7.

Dell'Angelica, E. C., Klumperman, J., Stoorvogel, W. and Bonifacino, J. S. (1998) Association of the AP-3 adaptor complex with clathrin. *Science*, **280**, 431-4.

Dick, T. P., Nussbaum, A. K., Deeg, M., Heinemeyer, W., Groll, M., Schirle, M., Keilholz, W., Stevanovic, S., Wolf, D. H., Huber, R., Rammensee, H. G. and Schild, H. (1998) Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. *J Biol Chem*, **273**, 25637-46.

Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H. and Groll, M. (1998) Conformational constraints for protein self-cleavage in the proteasome. *J Mol Biol*, **279**, 1187-91.

Dobson, C. M. (1992) Unfolded proteins, compact states and molten globules. *Curr Opin Struct Biol*, **2**, 6-12.

Du, W. and Risley, J. M. (2003) Acylation is rate-limiting in glycosylasparaginase-catalyzed hydrolysis of N4-(4'-substituted phenyl)-L-asparagines. *Org Biomol Chem*, **1**, 1900-5.

Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, E. J., Dodson, G. and Moody, P. C. E. (1995) Penicillin acylase has a single-amino-acid catalytic centre. *Nature*, **373**, 264-8.

Eckart, M. R. and Bussineau, C. M. (1996) Quality and authenticity of heterologous proteins synthesized in yeast. *Curr Opin Biotechnol*, **7**, 525-30.

Ellgaard, L. and Helenius, A. (2003) Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol*, **4**, 181-91.

Enomaa, N., Heiskanen, T., Halila, R., Sormunen, R., Seppälä, R., Vihinen, M. and Peltonen, L. (1992) Human aspartylglucosaminidase. A biochemical and immunochemical characterization of the enzyme in normal and aspartylglucosaminuria fibroblasts. *Biochem J*, **286**, 613-8.

Enomaa, N., Danos, O., Peltonen, L. and Jalanko, A. (1995) Correction of deficient enzyme activity in a lysosomal storage disease, aspartylglucosaminuria, by enzyme replacement and retroviral gene transfer. *Hum Gene Ther*, **6**, 723-31.

Farr, G. W., Scharl, E. C., Schumacher, R. J., Sondek, S. and Horwich, A. L. (1997) Chaperonin-mediated folding in the eukaryotic cytosol proceeds through rounds of release of native and nonnative forms. *Cell*, **89**, 927-37.

Fedorov, A. N. and Baldwin, T. O. (1997) GroE modulates kinetic partitioning of folding intermediates between alternative states to maximize the yield of biologically active protein. *J Mol Biol*, **268**, 712-23.

Fedorov, A. N. and Baldwin, T. O. (1998) Protein folding and assembly in a cell-free expression system. *Methods Enzymol*, **290**, 1-17.

Feng, Y., Sligar, S. G. and Wand, A. J. (1994) Solution structure of apocytochrome b562. *Nat Struct Biol*, **1**, 30-5.

Finley, D. and Chau, V. (1991) Ubiquitination. *Annu Rev Cell Biol*, **7**, 25-69.

Fisher, K. J., Tollersrud, O. K. and Aronson, N. N., Jr. (1990) Cloning and sequence analysis of a cDNA for human glycosylasparaginase. A single gene encodes the subunits of this lysosomal amidase. *FEBS Lett*, **269**, 440-4.

Fisher, K. J. and Aronson, N. N., Jr. (1991) Deletion of exon 8 causes glycosylasparaginase deficiency in an African American aspartylglucosaminuria (AGU) patient. *FEBS Lett*, **288**, 173-8.

Fisher, K. J., Klein, M., Park, H., Vettese, M. B. and Aronson, N. N., Jr. (1993) Post-translational processing and Thr-206 are required for glycosylasparaginase activity. *FEBS Lett*, **323**, 271-5.

Friant, S., Meier, K. D. and Riezman, H. (2003) Increased ubiquitin-dependent degradation can replace the essential requirement for heat shock protein induction. *Embo J*, **22**, 3783-91.

Fritz-Wolf, K., Koller, K. P., Lange, G., Liesum, A., Sauber, K., Schreuder, H., Aretz, W. and Kabsch, W. (2002) Structure-based prediction of modifications in glutarylamidase to allow single-step enzymatic production of 7-aminocephalosporanic acid from cephalosporin C. *Protein Sci*, **11**, 92-103.

Gallagher, T., Gilliland, G., Wang, L. and Bryan, P. (1995) The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure*, **3**, 907-14.

Gething, M. J. and Sambrook, J. (1992) Protein folding in the cell. *Nature*, **355**, 33-45.

Goodman, L. J. and Gorman, C. M. (1994) Autoproteolytic Activation of the Mouse Prohormone Convertase mPC1. *Biochem Biophys Res Commun*, **201**, 795-804.

Grallert, H. and Buchner, J. (2001) Review: a structural view of the GroE chaperone cycle. *J Struct Biol*, **135**, 95-103.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D. and Huber, R. (1997) Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature*, **386**, 463-71.

Grön, K., Aula, P. and Peltonen, L. (1990) Linkage of aspartylglucosaminuria (AGU) to marker loci on the long arm of chromosome 4. *Hum Genet*, **85**, 233-6.

Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L. and Comb, D. (1996) Activation of glycosylasparaginase: Formation of active N-terminal threonine by intramolecular autoproteolysis. *J Biol Chem*, **271**, 1732-7.

Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A., Whitaker, R. and Paulus, H. (1998) Characterization and functional analysis of the cis-autoproteolysis active center of glycosylasparaginase. *J Biol Chem*, **273**, 9695-702.

Guo, H. C., Xu, Q., Buckley, D. and Guan, C. (1998) Crystal structures of *Flavobacterium* glycosylasparaginase. An N- terminal nucleophile hydrolase activated by intramolecular proteolysis. *J Biol Chem*, **273**, 20205-12.

Haeuw, J. F., Grard, T., Alonso, C., Strecker, G. and Michalski, J. C. (1994) The core-specific lysosomal α(1-6)-mannosidase activity depends on aspartamidohydrolase activity. *Biochem J*, **297**, 463-6.

Halila, R., Baumann, M., Ikonen, E., Enomaa, N. and Peltonen, L. (1991) Human leucocyte aspartylglucosaminidase: Evidence for two different subunits in a more complex native structure. *Biochem J*, **276**, 251-6.

Hall, T. M., Porter, J. A., Young, K. E., Koonin, E. V., Beachy, P. A. and Leahy, D. J. (1997) Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell*, **91**, 85-97.

Haltia, M., Palo, J. and Autio, S. (1975) Aspartylglycosaminuria: a generalized storage disease. Morphological and histochemical studies. *Acta Neuropathol*, **31**, 243-55.

Hammond, C. and Helenius, A. (1995) Quality control in the secretory pathway. *Curr Opin Cell Biol*, **7**, 523-9.

Hancock, M. K., Haskins, D. J., Sun, G. and Dahms, N. M. (2002) Identification of residues essential for carbohydrate recognition by the insulin-like growth factor II/mannose 6-phosphate receptor. *J Biol Chem*, **277**, 11255-64.

Harkke, S., Laine, M. and Jalanko, A. (2003) Aspartylglucosaminidase (AGA) is efficiently produced and endocytosed by glial cells: implication for the therapy of a lysosomal storage disorder. *J Gene Med*, **5**, 472-82.

Hasilik, A., Waheed, A. and von Figura, K. (1981) Enzymatic phosphorylation of lysosomal enzymes in the presence of UDP-N-acetylglucosamine: Absence of activity in I-cell fibroblasts. *Biochem Biophys Res Commun*, **98**, 761-7.

Hebert, D. N., Foellmer, B. and Helenius, A. (1995) Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell*, **81**, 425-33.

Heiskanen, T., Tollersrud, O. K., Zhao, M. and Peltonen, L. (1994) Large-scale purification of human aspartylglucosaminidase: utilization of exceptional sodium dodecyl sulfate resistance. *Protein Expr Purif*, **5**, 205-10.

Helenius, A., Mellman, I., Wall, D. and Hubbard, A. L. (1983) Endosomes. *Trends Biochem Chem*, **8**, 245-50.

Hewitt, L., Kasche, V., Lummer, K., Lewis, R. J., Murshudov, G. N., Verma, C. S., Dodson, G. G. and Wilson, K. S. (2000) Structure of a slow processing precursor penicillin acylase from *Escherichia coli* reveals the linker peptide blocking the active-site cleft. *J Mol Biol*, **302**, 887-98.

Hicke, L. (2001) Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol*, **2**, 195-201.

Hiller, M. M., Finger, A., Schweiger, M. and Wolf, D. H. (1996) ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science*, **273**, 1725-8.

Hoflack, B. and Kornfeld, S. (1985) Lysosomal enzyme binding to mouse P388D1 macrophage membranes lacking the 215-kDa mannose 6-phosphate receptor: evidence for the existence of a second mannose 6-phosphate receptor. *Proc Natl Acad Sci U S A*, **82**, 4428-32.

Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A. and Nagata, K. (2001) A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. *EMBO Rep*, **2**, 415-22.

Hurtley, S. M., Bole, D. G., Hoover-Litty, H., Helenius, A. and Copeland, C. S. (1989) Interactions of misfolded influenza virus hemagglutinin with binding protein (BiP). *J Cell Biol*, **108**, 2117-26.

Ikonen, E., Aula, P., Grön, K., Tollersrud, O. K., Halila, R., Manninen, T., Syvänen, A.-C. and Peltonen, L. (1991a) Spectrum of mutations in aspartylglucosaminuria. *Proc Natl Acad Sci USA*, **88**, 11222-6.

Ikonen, E., Baumann, M., Grön, K., Syvänen, A. C., Enomaa, N., Halila, R., Aula, P. and Peltonen, L. (1991b) Aspartylglucosaminuria: cDNA encoding human aspartylglucosaminidase and the missense mutation causing the disease. *EMBO J*, **10**, 51-8.

Ikonen, E., Enomaa, N., Ulmanen, I. and Peltonen, L. (1991c) In vitro mutagenesis helps to unravel the biological consequences of aspartylglucosaminuria mutation. *Genomics*, **11**, 206-11.

Ikonen, E., Ulmanen, I. and Peltonen, L. (1992) Deletion of the 3'-untranslated region of aspartylglucosaminidase mRNA results in a lysosomal accumulation disease. *J Biol Chem*, **267**, 8715-8.

Ikonen, E., Julkunen, I., Tollersrud, O. K., Kalkkinen, N. and Peltonen, L. (1993) Lysosomal aspartylglucosaminidase is processed to the active subunit complex in the endoplasmic reticulum. *EMBO J*, **12**, 295-302.

Isoniemi, A., Hietala, M., Aula, P., Jalanko, A. and Peltonen, L. (1995) Identification of a novel mutation causing aspartylglucosaminuria reveals a mutation hotspot region in the aspartylglucosaminidase gene. *Hum Mut*, **5**, 318-26.

Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denisot, M. A., Badet, B., Polikarpov, I., Littlechild, J. A. and Teplyakov, A. (1996) Substrate binding is required for assembly of the active conformation of the catalytic site in Ntn amidotransferases: evidence from the 1.8 Å crystal structure of the glutaminase domain of glucosamine 6-phosphate synthase. *Structure*, **4**, 801-10.

Jakob, C. A., Burda, P., Roth, J. and Aebi, M. (1998) Degradation of misfolded endoplasmic reticulum glycoproteins in *Saccharomyces cerevisiae* is determined by a specific oligosaccharide structure. *J Cell Biol*, **142**, 1223-33.

Jalanko, A., Manninen, T. and Peltonen, L. (1995) Deletion of the C-terminal end of aspartylglucosaminidase resulting in a lysosomal accumulation disease: evidence for a unique genomic rearragement. *Hum Mol Genet*, **4**, 435-41.

Jalanko, A., Tenhunen, K., McKinney, C. E., LaMarca, M. E., Rapola, J., Autti, T., Joensuu, R., Manninen, T., Sipila, I., Ikonen, S., Riekkinen, P., Jr., Ginns, E. I. and Peltonen, L. (1998) Mice with an aspartylglucosaminuria mutation similar to humans replicate the pathophysiology in patients. *Hum Mol Genet*, **7**, 265-72.

Kaartinen, V. and Mononen, I. (1989) Analysis of aspartylglucosamine at the picomole level by high-performance liquid chromatography. *J Chromatogr*, **490**, 293-9.

Kaartinen, V. (1991) Glycoasparaginase in human urine. *Biochim Biophys Acta*, **1097**, 28-30.

Kaartinen, V., Williams, J. C., Tomich, J., Yates, J. R. I., Hood, L. E. and Mononen, I. (1991) Glycoasparaginase from human leukocytes. Inactivation and covalent modification with diazo-oxonorvaline. *J Biol Chem*, **266**, 5860-9.

Kaartinen, V., Mononen, T., Laatikainen, R. and Mononen, I. (1992) Substrate specificity and reaction mechanism of human glycoasparaginase. The N-glycosidic linkage of various glycoasparagines is cleaved through a reaction mechanism similar to L-asparaginase. *J Biol Chem*, **267**, 6855-8.

Kaartinen, V., Mononen, I., Voncken, J. W., Noronkoski, T., Gonzalez-Gomez, I., Heisterkamp, N. and Groffen, J. (1996) A mouse model for the human lysosomal disease aspartylglycosaminuria. *Nat Med*, **2**, 1375-8.

Kaartinen, V., Mononen, I., Gonzalez-Gomez, I., Noronkoski, T., Heisterkamp, N. and Groffen, J. (1998) Phenotypic characterization of mice with targeted disruption of glycosylasparaginase gene: a mouse model for aspartylglycosaminuria. *J Inherit Metab Dis*, **21**, 207-9.

Kasche, V., Lummer, K., Nurk, A., Piotraschke, E., Rieks, A., Stoeva, S. and Voelter, W. (1999) Intramolecular autoproteolysis initiates the maturation of penicillin amidase from *Escherichia coli*. *Biochim Biophys Acta*, **1433**, 76-86.

Katayama, T., Imaizumi, K., Sato, N., Miyoshi, K., Kudo, T., Hitomi, J., Morihara, T., Yoneda, T., Gomi, F., Mori, Y., Nakano, Y., Takeda, J., Tsuda, T., Itoyama, Y., Murayama, O., Takashima, A., St George-Hyslop, P., Takeda, M. and Tohyama, M. (1999) Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nat Cell Biol*, **1**, 479-85.

Kaushal, S., Ridge, K. D. and Khorana, H. G. (1994) Structure and function in rhodopsin: the role of asparagine-linked glycosylation. *Proc Natl Acad Sci U S A*, **91**, 4024-8.

Kiess, W., Blickenstaff, G. D., Sklar, M. M., Thomas, C. L., Nissley, S. P. and Sahagian, G. G. (1988) Biochemical evidence that the type II insulin-like growth factor receptor is identical to the cation-independent mannose 6-phosphate receptor. *J Biol Chem*, **263**, 9339-44.

Kim, J., Klein, P. G. and Mullet, J. E. (1991) Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. *J Biol Chem*, **266**, 14931-8.

Kim, J. H., Krahn, J. M., Tomchick, D. R., Smith, J. L. and Zalkin, H. (1996) Structure and function of the glutamine phosphoribosylpyrophosphate amidotransferase glutamine site and communication with the phosphoribosylpyrophosphate site. *J Biol Chem*, **271**, 15549-57.

Kim, J. K., Yang, I. S., Rhee, S., Dauter, Z., Lee, Y. S., Park, S. S. and Kim, K. H. (2003) Crystal Structures of Glutaryl 7-Aminocephalosporanic Acid Acylase: Insight into Autoproteolytic Activation. *Biochemistry*, **42**, 4084-93.

Kim, S. and Kim, Y. (2001) Active site residues of cephalosporin acylase are critical not only for enzymatic catalysis but also for post-translational modification. *J Biol Chem*, **276**, 48376-81.

Kim, Y., Yoon, K., Khang, Y., Turley, S. and Hol, W. G. (2000) The 2.0 Å crystal structure of cephalosporin acylase. *Structure Fold Des*, **8**, 1059-68.

Kim, Y., Kim, S., Earnest, T. N. and Hol, W. G. (2002) Precursor structure of cephalosporin acylase. Insights into autoproteolytic activation in a new N-terminal hydrolase family. *J Biol Chem*, **277**, 2823-9.

Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R. J. and Sacchettini, J. C. (1998) Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nature Struct Biol*, **5**, 31-6.

Klappa, P., Freedman, R. B. and Zimmermann, R. (1995) Protein disulphide isomerase and a lumenal cyclophilin-type peptidyl prolyl cis-trans isomerase are in transient contact with secretory proteins during late stages of translocation. *Eur J Biochem*, **232**, 755-64.

Kohno, M. and Yamashina, I. (1972) Purification and properties of 4-L-aspartyllycosamine amidohydrolase from hog kidney. *Biochem Biophys Acta*, **258**, 600-17.

Kopito, R. R. (1997) ER quality control: the cytoplasmic connection. *Cell*, **88**, 427-30.

Kornfeld, R. and Kornfeld, S. (1976) Comparative aspects of glycoprotein structure. *Annu Rev Biochem*, **45**, 217-37.

Kornfeld, R. and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem*, **54**, 631-64.

Kornfeld, S. and Mellman, I. (1989) The biogenesis of lysosomes. *Annu Rev Cell Biol*, **5**, 483-525.

Kuwajima, K. (1996) The molten globule state of alpha-lactalbumin. *FASEB J*, **10**, 102-9.

Kyttälä, A., Heinonen, O., Peltonen, L. and Jalanko, A. (1998) Expression and endocytosis of lysosomal aspartylglucosaminidase in mouse primary neurons. *J Neurosci*, **18**, 7750-6.

Laine, M., Richter, J., Fahlman, C., Rapola, J., Renlund, M., Peltonen, L., Karlsson, S. and Jalanko, A. (1999) Correction of peripheral lysosomal accumulation in mice with aspartylglucosaminuria by bone marrow transplantation. *Exp Hematol*, **27**, 1467-74.

Laitinen, A., Hietala, M., Haworth, J. C., Schroeder, M. L., Seargeant, L. E., Greenberg, C. R. and Aula, P. (1997) Two novel mutations in a Canadian family with aspartylglucosaminuria and early outcome post bone marrow transplantation. *Clin Genet*, **51**, 174-8.

Le Borgne, R. and Hoflack, B. (1998) Protein transport from the secretory to the endocytic pathway in mammalian cells. *Biochim Biophys Acta*, **1404**, 195-209.

Lee, J. L., Ekker, S. C., von Kessler, D. P., Porter, J. A., Sun, B. I. and Beachy, P. A. (1994) Autoproteolysis in *hedgehog* protein biogenesis. *Science*, **266**, 1528-37.

Lee, Y. S. and Park, S. S. (1998) Two-step autocatalytic processing of the glutaryl 7-aminocephalosporanic acid acylase from Pseudomonas sp. strain GK16. *J Bacteriol*, **180**, 4576-82.

Lee, Y. S., Kim, H. W. and Park, S. S. (2000) The role of alpha-amino group of the N-terminal serine of beta subunit for enzyme catalysis and autoproteolytic activation of glutaryl 7-aminocephalosporanic acid acylase. *J Biol Chem*, **275**, 39200-6.

Li, S., Smith, J. L. and Zalkin, H. (1999) Mutational analysis of *Bacillus subtilis* glutamine phosphoribosylpyrophosphate amidotransferase propeptide processing. *J Bacteriol*, **181**, 1403-8.

Liu, Y., Dunn, G. S. and Aronson, N. N., Jr. (1996) Purification, biochemistry and molecular cloning of an insect glycosylasparaginase from *Spodoptera frugiperda. Glycobiology*, **6**, 527-36.

Liu, Y., Guan, C. and Aronson, N. N., Jr. (1998) Site-directed mutagenesis of essential residues involved in the mechanism of bacterial glycosylasparaginase. *J Biol Chem*, **273**, 9688-94.

Liu, Y., Choudhury, P., Cabral, C. M. and Sifers, R. N. (1999) Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome. *J Biol Chem*, **274**, 5861-7.

Lough, T. J., Reddington, B. D., Grant, M. R., Hill, D. F., Reynolds, P. H. and Farnden, K. J. (1992) The isolation and characterisation of a cDNA clone encoding L-asparaginase from developing seeds of lupin (*Lupinus arboreus*). *Plant Mol Biol*, **19**, 391-9.

Ludwig, T., Munier-Lehmann, H., Bauer, U., Hollinshead, M., Ovitt, C., Lobel, P. and Hoflack, B. (1994) Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts. *Embo J*, **13**, 3430-7.

Luzio, J. P., Rous, B. A., Bright, N. A., Pryor, P. R., Mullock, B. M. and Piper, R. C. (2000) Lysosome-endosome fusion and lysosome biogenesis. *J Cell Sci*, **113**, 1515-24.

Lönnqvist, L., Karttunen, L., Rantamäki, T., Kielty, C., Raghunath, M. and Peltonen, L. (1996) A point mutation creating an extra N-glycosylation site in fibrillin-1 results in neonatal Marfan syndrome. *Genomics*, **36**, 468-75.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science*, **268**, 533-9.

Mahadevan, S. and Tappel, A. L. (1967) Beta-aspartylglucosylamine amido hydrolase of rat liver and kidney. *J Biol Chem*, **242**, 4568-76.

Makino, M., Kojima, T. and Yamashina, I. (1966) Enzymatic cleavage of glycopeptides. *Biochem Biophys Res Commun*, **24**, 961-6.

Makino, M., Kojima, T., Ohgushi, T. and Yamashina, I. (1968) Studies on enzymes acting on glycopeptides. *J Biochem*, **63**, 186-92.

Mann, M. and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat Biotechnol*, **21**, 255-61.

Marshall, R. D. (1974) The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochem Soc Symp*, 17-26.

Matthews, B. W., Sigler, P. B., Henderson, R. and Blow, D. M. (1967) Three-dimensional structure of tosyl-alpha-chymotrypsin. *Nature*, **214**, 652-6.

Maury, C. P. (1980) Accumulation of glycoprotein-derived metabolites in neural and visceral tissue in aspartylglycosaminuria. *J Lab Clin Med*, **96**, 838-44.

McCann, P. J., III, O'Boyle, D. R. and Deckman, I. C. (1994) Investigation of the specificity of the herpes simplex virus type 1 protease by point mutagenesis of the autoproteolysis sites. *J Virol*, **68**, 526-9.

McCracken, A. A. and Brodsky, J. L. (1996) Assembly of ER-associated protein degradation in vitro: dependence on cytosol, calnexin, and ATP. *J Cell Biol*, **132**, 291-8.

McVey, C. E., Walsh, M. A., Dodson, G. G., Wilson, K. S. and Brannigan, J. A. (2001) Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. *J Mol Biol*, **313**, 139-50.

Mellman, I., Fuchs, R. and Helenius, A. (1986) Acidification of the endocytic and exocytic pathways. *Annu Rev Biochem*, **55**, 663-700.

Miele, R. G., Castellino, F. J. and Bretthauer, R. K. (1997a) Characterization of the acidic oligosaccharides assembled on the *Pichia pastoris*-expressed recombinant kringle 2 domain of human tissue-type plasminogen activator. *Biotechnol Appl Biochem*, **26**, 79-83.

Miele, R. G., Nilsen, S. L., Brito, T., Bretthauer, R. K. and Castellino, F. J. (1997b) Glycosylation properties of the *Pichia pastoris*-expressed recombinant kringle 2 domain of tissue-type plasminogen activator. *Biotechnol Appl Biochem*, **25**, 151-7.

Milner-White, E. J. (1990) Situations of gamma-turns in proteins. Their relation to alpha-helices, beta-sheets and ligand binding sites. *J Mol Biol*, **216**, 386-97.

Mononen, I., Fisher, K. J., Kaartinen, V. and Aronson, N. N., Jr. (1993) Aspartylglucosaminuria: protein chemistry and molecular biology of the most common lysosomal storage disease of glycoprotein degradation. *FASEB J*, **7**, 1247-56.

Morris, C., Heisterkamp, N., Groffen, J., Williams, J. C. and Mononen, I. (1992) Chromosomal localization of the human glycoasparaginase gene to 4q32-q33. *Hum Genet*, **88**, 295-7.

Mortimore, G. E., Lardeux, B. R. and Adams, C. E. (1988) Regulation of microautophagy and basal protein turnover in rat liver. Effects of short-term starvation. *J Biol Chem*, **263**, 2506-12.

Muchmore, C. R., Krahn, J. M., Kim, J. H., Zalkin, H. and Smith, J. L. (1998) Crystal structure of glutamine phosphoribosylpyrophosphate amidotransferase from *Escherichia coli*. *Protein Sci*, **7**, 39-51.

Mueller, O. T., Honey, N. K., Little, L. E., Miller, A. L. and Shows, T. B. (1983) Mucolipidosis II and III. The genetic relationships between two disorders of lysosomal enzyme biosynthesis. *J Clin Invest*, **72**, 1016-23.

Murray, P. J., Watowich, S. S., Lodish, H. F., Young, R. A. and Hilton, D. J. (1995) Epitope tagging of the human endoplasmic reticulum HSP70 protein, BiP, to facilitate analysis of BiP--substrate interactions. *Anal Biochem*, **229**, 170-9.

Nilssen, O., Tollersrud, O. K., Borud, O. and Tranebjaerg, L. (1993) A simple and rapid PCR based method for AGU(Fin) determination. *Hum Mol Genet*, **2**, 484.

Nilssen, O., Berg, T., Riise, H. M., Ramachandran, U., Evjen, G., Hansen, G. M., Malm, D., Tranebjaerg, L. and Tollersrud, O. K. (1997) alpha-Mannosidosis: functional cloning of the lysosomal alpha-mannosidase cDNA and identification of a mutation in two affected siblings. *Hum Mol Genet*, **6**, 717-26.

Norio, R., Nevanlinna, H. R. and Perheentupa, J. (1973) Hereditary diseases in Finland; rare flora in rare soul. *Ann Clin Res*, **5**, 109-41.

Norio, R. (2003a) Finnish Disease Heritage I: characteristics, causes, background. *Hum Genet*, **112**, 441-56.

Norio, R. (2003b) Finnish Disease Heritage II: population prehistory and genetic roots of Finns. *Hum Genet*, **112**, 457-69.

Norio, R. (2003c) The Finnish Disease Heritage III: the individual diseases. *Hum Genet*, **112**, 470-526.

Noronkoski, T. and Mononen, I. (1997) Influence of L-fucose attached alpha 1-->6 to the asparagine-linked N-acetylglucosamine on the hydrolysis of the N-glycosidic linkage by human glycosylasparaginase. *Glycobiology*, **7**, 217-20.

Noronkoski, T., Stoineva, I. B., Petkov, D. D. and Mononen, I. (1997) Recombinant human glycosylasparaginase catalyzes hydrolysis of L-asparagine. *FEBS Lett*, **412**, 149-52.

Noronkoski, T., Stoineva, I. B., Ivanov, I. P., Petkov, D. D. and Mononen, I. (1998) Glycosylasparaginase-catalyzed synthesis and hydrolysis of beta-aspartyl peptides. *J Biol Chem*, **273**, 26295-7.

Novikoff, A. (1963) Lysosomes in the physiology and pathology of cells: Contributions of staining methods. In *Lysosomes*, 151-75, Ciba Found. Symposium, London

Oinonen, C., Tikkanen, R., Rouvinen, J. and Peltonen, L. (1995) Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nat Struct Biol*, **2**, 1102-8.

Oinonen, C. and Rouvinen, J. (2000) Structural comparison of Ntn-hydrolases. *Protein Sci*, **9**, 2329-37.

Olson, L. J., Zhang, J., Dahms, N. M. and Kim, J. J. (2002) Twists and turns of the cation-dependent mannose 6-phosphate receptor. Ligand-bound versus ligand-free receptor. *J Biol Chem*, **277**, 10156-61.

Palo, J. and Mattsson, K. (1970) Eleven new cases of aspartylglucosaminuria. *J Ment Defic Res*, **14**, 168-73.

Palo, J., Riekkinen, P., Arstila, A. U., Autio, S. and Kivimaki, T. (1972) Aspartylglucosaminuria. II. Biochemical studies on brain, liver, kidney and spleen. *Acta Neuropathol*, **20**, 217-24.

Park, H., Fisher, K. J. and Jr., A. N. N. (1991) Genomic structure of human lysosomal glycosylasparaginase. *FEBS Lett*, **288**, 168-72.

Park, H., Vettese, M. B., Fensom, A. H., Fisher, K. J. and Aronson, N. N., Jr. (1993) Characterization of three alleles causing aspartylglucosaminuria. *Biochem J*, **290**, 735-41.

Park, H., Rossiter, M., Fensom, A. H., Winchester, B. and Aronson, N. N., Jr. (1996a) Single base deletion in exon 7 of the glycoasparaginase gene causes a mild form of aspartylglucosaminuria in a patient of Mauritian origin. *J Inherit Metab Dis*, **19**, 76-83.

Park, H., Vettese-Dadey, M. and Aronson, N. N., Jr. (1996b) Glycosylation and phosphorylation of lysosomal glycosylasparaginase. *Arch Biochem Biophys*, **328**, 73-7.

Pastinen, T., Perola, M., Ignatius, J., Sabatti, C., Tainola, P., Levander, M., Syvänen, A. C. and Peltonen, L. (2001) Dissecting a population genome for targeted screening of disease mutations. *Hum Mol Genet*, **10**, 2961-72.

Pearse, B. M. and Robinson, M. S. (1990) Clathrin, adaptors, and sorting. *Annu Rev Cell Biol*, **6**, 151-71.

Peltola, M., Chitayat, D., Peltonen, L. and Jalanko, A. (1994) Characterization of a point mutation in aspartylglucosaminidase gene: evidence for a readthrough of a translational stop codon. *Hum Mol Genet*, **3**, 2237-42.

Peltola, M., Tikkanen, R., Peltonen, L. and Jalanko, A. (1996) Ser72Pro active-site disease mutation in human lysosomal aspartylglucosaminidase: Abnormal intracellular processing and evidence for extracellular activation. *Hum Mol Genet*, **5**, 737-43.

Peltola, M., Kyttälä, A., Heinonen, O., Rapola, J., Paunio, T., Revah, F., Peltonen, L. and Jalanko, A. (1998) Adenovirus-mediated gene transfer results in decreased lysosomal storage in brain and total correction in liver of aspartylglucosaminuria (AGU) mouse. *Gene Ther*, **5**, 1314-21.

Peltonen, L., Jalanko, A. and Varilo, T. (1999) Molecular genetics of the Finnish disease heritage. *Hum Mol Genet*, **8**, 1913-23.

Perheentupa, J. (1972) Suomalainen tautiperintö (Symposium on inherited diseases in Finland). *Duodecim*, **88**, 1-166.

Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J. and Belfort, M. (1994) Protein splicing elements: inteins and exteins--a definition of terms and recommended nomenclature. *Nucleic Acids Res*, **22**, 1125-7.

Perler, F. B., Olsen, G. J. and Adam, E. (1997) Compilation and analysis of intein sequences. *Nucleic Acids Res*, **25**, 1087-93.

Petrovan, R. J., Govers-Riemslag, J. W. P., Nowak, G., Hemker, H. C., Tans, G. and Rosing, J. (1998) Autocatalytic Peptide Bond Cleavages in Prothrombin and Meizothrombin. *Biochemistry*, **37**, 1185-91.

Poland, B. W., Xu, M. Q. and Quiocho, F. A. (2000) Structural insights into the protein splicing mechanism of PI-SceI. *J Biol Chem*, **275**, 16408-13.

Pollitt, R. J., Jenner, F. A. and Merskey, H. (1968) Aspartylglucosaminuria. An inborn error of metabolism associated with mental defect. *Lancet*, **2**, 253-5.

Presley, J. F., Cole, N. B., Schroer, T. A., Hirschberg, K., Zaal, K. J. and Lippincott-Schwartz, J. (1997) ER-to-Golgi transport visualized in living cells. *Nature*, **389**, 81-5.

Qian, X., Guan, C. and Guo, H. C. (2003) A dual role for an aspartic acid in glycosylasparaginase autoproteolysis. *Structure*, **11**, 997-1003.

Quentin, E., Gladen, A., Roden, L. and Kresse, H. (1990) A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc Natl Acad Sci U S A*, **87**, 1342-6.

Raas-Rothschild, A., Cormier-Daire, V., Bao, M., Genin, E., Salomon, R., Brewer, K., Zeigler, M., Mandel, H., Toth, S., Roe, B., Munnich, A. and Canfield, W. M. (2000) Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). *J Clin Invest*, **105**, 673-81.

Reissig, J. L., Strominger, J. L. and Leloir, K. M. (1955) A modified colorimetric method for the estimation of N-acetylamino sugars. *J Biol Chem*, **217**, 959-66.

Reitman, M. L. and Kornfeld, S. (1981) Lysosomal enzyme targeting. *N*-acetylglucosaminylphosphotransferase selectively phosphorylates native lysosomal enzymes. *J Biol Chem*, **256**, 11977-80.

Reitman, M. L., Varki, A. and Kornfeld, S. (1981) Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity. *J Clin Invest*, **67**, 1574-9.

Richards, F. M. (1974) The interpretation of protein structures: total volume, group volume distributions and packing density. *J Mol Biol*, **82**, 1-14.

Riikonen, A., Ikonen, E., Sormunen, R., Lehto, V.-P., Peltonen, L. and Jalanko, A. (1994) Dissection of the molecular consequences of a double mutation causing a human lysosomal disease. *DNA Cell Biol*, **13**, 257-64.

Riikonen, A., Tikkanen, R., Jalanko, A. and Peltonen, L. (1995) Immediate interaction between the nascent subunits and two conserved amino acids Trp^{34} and Thr^{206} are needed for the catalytic activity of aspartylglucosaminidase. *J Biol Chem*, **270**, 4903-7.

Riikonen, A., Rouvinen, J., Tikkanen, R., Julkunen, I., Peltonen, L. and Jalanko, A. (1996) Primary folding of aspartylglucosaminidase. Significance of disulfide bridges and evidence of early multimerization. *J Biol Chem*, **271**, 21340-4.

Rip, J. W., Coulter-Mackie, M. B., Rupar, C. A. and Gordon, B. A. (1992) Purification and structure of human liver aspartylglucosaminidase. *Biochem J*, **288**, 1005-10.

Risley, J. M., Huang, D. H., Kaylor, J. J., Malik, J. J. and Xia, Y. Q. (2001a) Glycosylasparaginase inhibition studies: competitive inhibitors, transition state mimics, noncompetitive inhibitors. *J Enzyme Inhib*, **16**, 269-74.

Risley, J. M., Huang, D. H., Kaylor, J. J., Malik, J. J., Xia, Y. Q. and York, W. M. (2001b) Glycosylasparaginase activity requires the alpha-carboxyl group, but not the alpha-amino group, on N(4)-(2-Acetamido-2-deoxy-beta-D- glucopyranosyl)-L-asparagine. *Arch Biochem Biophys*, **391**, 165-70.

Rivera, V. M., Wang, X., Wardwell, S., Courage, N. L., Volchuk, A., Keenan, T., Holt, D. A., Gilman, M., Orci, L., Cerasoli, F., Jr., Rothman, J. E. and Clackson, T. (2000) Regulation of protein secretion through controlled aggregation in the endoplasmic reticulum. *Science*, **287**, 826-30.

Roberts, D. L., Weix, D. J., Dahms, N. M. and Kim, J. J. (1998) Molecular basis of lysosomal enzyme recognition: three-dimensional structure of the cation-dependent mannose 6-phosphate receptor. *Cell*, **93**, 639-48.

Robinson, M. S. (1994) The role of clathrin, adaptors and dynamin in endocytosis. *Curr Opin Cell Biol*, **6**, 538-44.

Roder, H. and Colon, W. (1997) Kinetic role of early intermediates in protein folding. *Curr Opin Struct Biol*, **7**, 15-28.

Rosorius, O., Issinger, O. G. and Braulke, T. (1993) Phosphorylation of the cytoplasmic tail of the 300-kDa mannose 6-phosphate receptor is required for the interaction with a cytosolic protein. *J Biol Chem*, **268**, 21470-3.

Roth, R. A. and Pierce, S. B. (1987) In vivo cross-linking of protein disulfide isomerase to immunoglobulins. *Biochemistry*, **26**, 4179-82.

Sahagian, G. G., Distler, J. and Jourdian, G. W. (1981) Characterization of a membrane-associated receptor from bovine liver that binds phosphomannosyl residues of bovine testicular β-galactosidase. *Proc Natl Acad Sci USA*, **78**, 4289-93.

Saier, M. H., Jr. (1999) Eukaryotic transmembrane solute transport systems. *Int Rev Cytol*, **190**, 61-136.

Saraste, J. and Kuismanen, E. (1984) Pre- and post-Golgi vacuoles operate in the transport of Semliki Forest virus membrane glycoproteins to the cell surface. *Cell*, **38**, 535-49.

Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frömmel, C., Löwe, J., Huber, R., Kloetzel, P. M. and Schmidt, M. (1996) Analysis of mammalian 20S proteasome biogenesis: the maturation of β-subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J*, **15**, 6887-98.

Schweizer, A., Kornfeld, S. and Rohrer, J. (1996) Cysteine34 of the cytoplasmic tail of the cation-dependent mannose 6-phosphate receptor is reversibly palmitoylated and required for normal trafficking and lysosomal enzyme sorting. *J Cell Biol*, **132**, 577-84.

Semenza, J. C., Hardwick, K. G., Dean, N. and Pelham, H. R. (1990) ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell*, **61**, 1349-57.

Shao, Y., Xu, M. Q. and Paulus, H. (1996) Protein splicing: evidence for an N-O acyl rearrangement as the initial step in the splicing process. *Biochemistry*, **35**, 3810-5.

Smith, J. L., Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H. and Satow, Y. (1994) Structure of the allosteric regulatory enzyme of purine biosynthesis. *Science*, **264**, 1427-33.

Stevens, T. H. and Forgac, M. (1997) Structure, function and regulation of the vacuolar (H+)-ATPase. *Annu Rev Cell Dev Biol*, **13**, 779-808.

Stinchcombe, J. C. and Griffiths, G. M. (1999) Regulated secretion from hemopoietic cells. *J Cell Biol*, **147**, 1-6.

Sugahara, K., Nishimura, K., Aula, P. and Yamashina, I. Y. (1976) Enzymatic determination of urinary aspartylglycosylamine: a rapid and sensitive-method to detect aspartylclycosylaminuria (AGU). *Clin Chim Acta*, **72**, 265-7.

Suresh, C. G., Pundle, A. V., SivaRaman, H., Rao, K. N., Brannigan, J. A., McVey, C. E., Verma, C. S., Dauter, Z., Dodson, E. J. and Dodson, G. G. (1999) Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. *Nat Struct Biol*, **6**, 414-6.

Syvänen, A. C., Ikonen, E., Manninen, T., Bengtsröm, M., Söderlund, H., Aula, P. and Peltonen, L. (1992) Convenient and quantitative determination of the frequency of a mutant allele using solid-phase minisequencing: Application to aspartylglucosaminuria in Finland. *Genomics*, **12**, 590-5.

Tarentino, A., Plummer, T. and Maley, F. (1975) The isolation and structure of the core oligosaccharide sequences of IgM. *Biochemistry*, **14**, 5516-23.

Tarentino, A. L. and Maley, F. (1969) The purification and properties of a β-aspartyl N-acetylglucosylamine amidohydrolase from hen oviduct. *Arch Biochem Biophys*, **130**, 295-303.

Tarentino, A. L. and Plummer, T. H., Jr. (1993) The first demonstration of a prokaryotic glycosylasparaginase. *Biochem Biophys Res Commun*, **197**, 179-86.

Tarentino, A. L., Quinones, G., Hauer, C. R., Changchien, L. M. and Plummer, T. H., Jr. (1995) Molecular cloning and sequence analysis of *Flavobacterium meningosepticum* glycosylasparaginase: a single gene encodes the alpha and beta subunits. *Arch Biochem Biophys*, **316**, 399-406.

Tenhunen, K., Laan, M., Manninen, T., Palotie, A., Peltonen, L. and Jalanko, A. (1995) Molecular cloning, chromosomal assingment and expression of the mouse aspartylglucosaminidase gene. *Genomics*, **30**, 244-50.

Tenhunen, K., Uusitalo, A., Autti, T., Joensuu, R., Kettunen, M., Kauppinen, R. A., Ikonen, S., LaMarca, M. E., Haltia, M., Ginns, E. I., Jalanko, A. and Peltonen, L. (1998) Monitoring the CNS pathology in aspartylglucosaminuria mice. *J Neuropathol Exp Neurol*, **57**, 1154-63.

Tikkanen, R., Enomaa, N., Riikonen , A., Ikonen, E. and Peltonen, L. (1995) Intracellular sorting of aspartylglucosaminidase: The role of N-linked oligosaccharides and evidence of Man-6-P independent lysosomal targeting. *DNA Cell Biol*, **14**, 305-12.

Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, R. and Peltonen, L. (1996a) Functional analyses of active site residues of human lysosomal aspartylglucosaminidase: Implications for catalytic mechanism and autocatalytic activation. *EMBO J*, **15**, 2954-60.

Tikkanen, R., Rouvinen, J., Torronen, A., Kalkkinen, N. and Peltonen, L. (1996b) Large-scale purification and preliminary x-ray diffraction studies of human aspartylglucosaminidase. *Proteins*, **24**, 253-8.

Tikkanen, R., Peltola, M., Oinonen, C., Rouvinen, J. and Peltonen, L. (1997) Several cooperating binding sites mediate the interaction of a lysosomal enzyme with phosphotransferase. *Embo J*, **16**, 6684-93.

Tollersrud, O. K., Hofmann, S. H. and Aronson, N. N., Jr. (1988) The turnover of lysosomal glycosylasparaginase in rat liver. *Biochim Biophys Acta*, **953**, 353-6.

Tollersrud, O. K. and Aronson, N. N., Jr. (1989) Purification and characterization of rat liver glycosylasparaginase. *Biochem J*, **260**, 101-8.

Tollersrud, O. K. and Aronson, N. N., Jr. (1992) Comparison of liver glycosylasparaginases from six vertebrates. *Biochem J*, **282**, 891-7.

Tollersrud, O. K., Heiskanen, T. and Peltonen, L. (1994a) Human leucocyte glycosylasparaginase is an alpha/beta-heterodimer of 19 kDa alpha-subunit and 17 and 18 kDa beta-subunit. *Biochem J*, **300**, 541-4.

Tollersrud, O. K., Nilssen, O., Tranebjaerg, L. and Borud, O. (1994b) Aspartylglucosaminuria in northern Norway: a molecular and genealogical study. *J Med Genet*, **31**, 360-3.

Uusitalo, A., Tenhunen, K., Tenhunen, J., Matikainen, S., Peltonen, L. and Jalanko, A. (1997) Expression and regulation of the human and mouse aspartylglucosaminidase gene. *J Biol Chem*, **272**, 9524-30.

Uusitalo, A., Tenhunen, K., Heinonen, O., Hiltunen, J. O., Saarma, M., Haltia, M., Jalanko, A. and Peltonen, L. (1999) Toward understanding the neuronal pathogenesis of aspartylglucosaminuria: expression of aspartylglucosaminidase in brain during development. *Mol Genet Metab*, **67**, 294-307.

Valkonen, S., Hietala, M., Savontaus, M. L. and Aula, P. (1999) Origin of Finnish mutations causing aspartylglucosaminuria. *Hereditas*, **131**, 191-5.

Varilo, T., Nikali, K., Suomalainen, A., Lönnqvist, T. and Peltonen, L. (1996) Tracing an ancestral mutation: genealogical and haplotype analysis of the infantile onset spinocerebellar ataxia locus. *Genome Res*, **6**, 870-5.

Vernet, T., Khouri, H. E., Laflamme, P., Tessier, D. C., Musil, R., Gour-Salin, B. J., Storer, A. C. and Thomas, D. Y. (1991) Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing. *J Biol Chem*, **266**, 21451-7.

Vogel, J. P., Misra, L. M. and Rose, M. D. (1990) Loss of BiP/GRP78 function blocks translocation of secretory proteins in yeast. *J Cell Biol*, **110**, 1885-95.

Waheed, A., Hasilik, A. and von Figura, K. (1981) Processing of the phosphorylated recognition marker in lysosomal enzymes. Characterization and partial purification of a microsomal α-*N*-acetylglucosaminyl phosphodiesterase. *J Biol Chem*, **256**, 5717-21.

Walkley, S. U. (1998) Cellular pathology of lysosomal storage disorders. *Brain Pathol*, **8**, 175-93.

Wang, A. M., Schindler, D. and Desnick, R. (1990) Schindler disease: the molecular lesion in the alpha-N-acetylgalactosaminidase gene that causes an infantile neuroaxonal dystrophy. *J Clin Invest*, **86**, 1752-6.

Wang, Y., Liang, Z. H., Zhang, Y. S., Yao, S. Y., Xu, Y. G., Tang, Y. H., Zhu, S. Q., Cui, D. F. and Feng, Y. M. (2001) Human insulin from a precursor overexpressed in the methylotrophic yeast *Pichia pastoris* and a simple procedure for purifying the expression product. *Biotechnol Bioeng*, **73**, 74-9.

Wang, Y. and Guo, H. C. (2003) Two-step Dimerization for Autoproteolysis to Activate Glycosylasparaginase. *J Biol Chem*, **278**, 3210-9.

Wang, Z. Q., Fung, M. R., Barlow, D. P. and Wagner, E. F. (1994) Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature*, **372**, 464-7.

Warshel, A., Naray-Szabo, G., Sussman, F. and Hwang, J. K. (1989) How do serine proteases really work? *Biochemistry*, **28**, 3629-37.

Wu, Z., Yao, N., Le, H. V. and Weber, P. C. (1998) Mechanism of autoproteolysis at the NS2-NS3 junction of the hepatitis C virus polyprotein. *Trends Biochem Sci*, **23**, 92-4.

Xu, Q., Buckley, D., Guan, C. and Guo, H. C. (1999) Structural insights into the mechanism of intramolecular proteolysis. *Cell*, **98**, 651-61.

Xuan, J., Tarentino, A. L., Grimwood, B. G., Plummer, T. H., Jr., Cui, T., Guan, C. and Van Roey, P. (1998) Crystal structure of glycosylasparaginase from *Flavobacterium meningosepticum*. *Protein Sci*, **7**, 774-81.

Yamashita, K., Tachibana, Y., Takada, S., Matsuda, I., Arashima, S. and Kobata, A. (1979) Urinary glycopeptides of fucosidosis. *J Biol Chem*, **254**, 4820-7.

Yoon, J., Oh, B., Kim, K., Park, J., Han, D., Kim, K. K., Cha, S. S., Lee, D. and Kim, Y. (2003) A bound water molecule is crucial in initiating auto-catalytic precursor activation in a N-terminal hydrolase. *J Biol Chem*.

Yoshida, K., Yanagisawa, N., Oshima, A., Sakuraba, H., Iida, Y. and Suzuki, Y. (1992) Splicing defect of the glycoasparaginase gene in two Japanese siblings with aspartylglucosaminuria. *Hum Genet*, **90**, 179-80.

Zerangue, N., Schwappach, B., Jan, Y. N. and Jan, L. Y. (1999) A new ER trafficking signal regulates the subunit stoichiometry of plasma membrane K(ATP) channels. *Neuron*, **22**, 537-48.

Zimmerman, U. J. and Schlaepfer, W. W. (1991) Two-stage autolysis of the catalytic subunit initiates activation of calpain I. *Biochim Biophys Acta*, **1078**, 192-8.

Zlotnick, A., Reddy, V. S., Dasgupta, R., Schneemann, A., Ray, W. J., Jr., Rueckert, R. R. and Johnson, J. E. (1994) Capsid assembly in a family of animal viruses primes an autoproteolytic maturation that depends on a single aspartic acid residue. *J Biol Chem*, **269**, 13680-4.